Commodity Futures Characteristics and Asset Pricing Models

July 1, 2024

Qin Yiyi*, Jun Cai, Jie Zhu, and Robert Webb

Abstract

A latent-factor model based on the Instrumented Principal Component Analysis (IPCA) methodology of Kelly et al. (2019) outperforms existing factor models in explaining cross-sectional variations in commodity futures returns. The model allows for observed commodity futures characteristics to work as instruments for unobservable dynamic factor loadings. We find that the relationship between characteristics and commodity futures returns is driven by compensation for exposure to latent risk factors (beta) rather than compensation for exposure to mispricing (alpha). Three latent factors deliver more powerful explanations than any number of observable factors. Among a collection of twenty characteristics, only three are significantly related to latent factor betas. These three characteristics are momentum, expected shortfall, and idiosyncratic volatility.

Keywords: Commodity futures contracts; Observable risk factors; Latent factor models; Instrumented principal component analysis

*Qin: corresponding author, yiyiqin@swufe.edu.cn, School of Finance, Southwestern University of Finance and Economics, Chengdu, Sichuan, P. R. China.

Cai: efjuncai@cityu.edu.hk, Department of Economics and Finance, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong, P. R. China.

Zhu: zhu_jie@t.shu.edu.cn, SILC Business School, Shanghai University, Shanghai, P. R. China.

Webb: riw4j@virginia.edu, McIntire School of Commerce, University of Virginia, Charlottesville, VA, U.S.A.

We benefited from discussions with Robert Korajczyk, Michael Melvin, Seth Pruitt, and Yinan Su regarding the IPCA and PPCA model estimation. We are also grateful to John Hua Fan for conversations regarding the S&P Goldman Sachs commodity total return indices.

1. Introduction

A large body of literature has established that commodity futures investment strategies based on fundamental commodity characteristics such as momentum, value, basis, basis momentum, hedging pressure, idiosyncratic volatility, and skewness earn significantly higher returns relative to commodity market indices (De Roon et al., 2000; Miffre and Rallis, 2007; Asness et al., 2013; Yang, 2013; Bianchi et al., 2015; Fernandez-Perez et al., 2018; Bakshi et al., 2019; Boons and Prado, 2019).[1] Typically, these studies sort commodity futures contracts based on one fundamental characteristic to generate hedge portfolio returns and then evaluate those hedge portfolio returns against alternative benchmarks. Significant alphas are interpreted as risk-adjusted abnormal returns. Bianchi et al. (2015) study trading strategies that jointly exploit momentum and contrarian returns in commodity futures. Fuertes et al. (2015) design a triple-screen trading strategy based on three signals: momentum, term-structure (roll yields), and idiosyncratic volatility. Fernandez-Perez et al. (2019) and Fuertes and Zhao (2023) investigate the issue of style integration, combining multiple characteristics to form a portfolio with *simultaneous* exposure to many styles including momentum, value, carry (roll yields), liquidity, skewness, and basis momentum.

The commodity risk premia literature, so far, has not addressed the sources of these abnormal returns, i.e., whether abnormal returns are due to alphas or betas related to sorting characteristics. This alpha versus beta debate has been a long-standing issue for the equity market since the work of Fama and French (1993). Daniel and Titman (1997) attempt to address this issue using portfolio sorting based on lagged beta estimates and firm characteristics. They find significant characteristic-based returns *controlling for betas* but not beta-based returns *controlling for characteristics*. However, the double-sorting approach can only handle one characteristic at a time. In addition, there exist some statistical issues (Ferson and Harvey, 1997; Berk, 2000; Kim et al., 2021).

Kelly et al. (2019) address the alpha versus beta issue by developing an instrumented principal component analysis (IPCA) methodology. Kelly et al.'s (2019) model assumes that the relationship between either alpha or beta and characteristics is constant over the full sample period (or full estimation period). They conclude that firm characteristics can predict the cross-section of stock returns *because betas*, rather than alphas, *are related to these characteristics*.

[1] Miffre (2016) provides a comprehensive review of various investment strategies in the commodity futures market. See also Sakkas and Tessaromatis (2020) for a brief summary.

This study attempts to disentangle the alpha versus beta effect in generating abnormal returns for commodity futures. We follow the IPCA approach of Kelly et al. (2019). The IPCA approach is suitable to study the commodities market for the following three reasons. First, the IPCA model allows characteristics to have short-term dynamics, i.e., characteristics change on a monthly basis. This is the case for commodity futures. All commodity futures characteristics, such as momentum, basis, hedging pressure, speculative pressure, and basis momentum, can be measured monthly.[2] Second, the IPCA approach can simultaneously include multiple traditional observable risk factors or pre-specified risk factors within an IPCA specification. This makes it easy to test the incremental contribution of each observable risk factor in the presence of latent risk factors. Third, the IPCA approach allows us to test for the significance of each individual characteristic after controlling for other characteristics in a complete multivariate analysis.

Our dataset covers 34 commodity futures contracts that fall into the following five categories: energy, grains and oilseeds, livestock, metals, and softs. Coverage is similar to that used in earlier studies. We consider a total of 20 commodity futures characteristics widely used in the literature: 3-month momentum, 12-month momentum, 18-month contrarian return, 36-month contrarian return, 52-week high, idiosyncratic volatility, skewness, maximum daily return, expected shortfall, basis, hedging pressure, speculative pressure, 3-month basis momentum, 12-month basis momentum, commodity market beta, U.S. dollar index beta, inflation beta, trading volume, open interest, and the Amihud liquidity measure. We also construct seven observable risk factors related to market, momentum, basis, hedging pressure and speculative pressure, basis momentum, idiosyncratic volatility, and skewness (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*).

Our major findings can be summarized as follows. First, the impact of characteristics is negligible in the alphas of latent factors. With three latent factors, we cannot reject the null hypothesis that alpha is zero. In other words, commodity futures characteristics do not generate abnormal returns through the mispricing channel. Among 20 characteristics, only three make a substantial contribution to explaining betas in latent factor models. When controlling for

[2] Alternatively, Kim et al. (2021) propose a projected principal component analysis (PPCA) approach to investigate the alpha versus beta debate. The PPCA rolling-window approach uses a set of characteristics that is fixed at the beginning of the estimation window (12 months, for example). Then the PPCA approach projects the monthly returns of N stocks during the 12-month period onto the set of characteristics. This requires a balanced panel of return data. The PPCA next extracts the eigenvectors from the projected returns. The consistency of the PPCA approach is established based on a large number of cross-sectional stocks (N) and other stationarity assumptions. The rolling-window approach is appropriate when characteristics include annual accounting ratios that are constant over the 12-month estimation period.

momentum, expected short fall, and idiosyncratic volatility, none of the remaining characteristics deliver significant contribution in explaining betas of latent factors. The three significant characteristics in the betas of latent factors remain robust even when we simultaneously control for observable factors in the IPCA model and allow betas of observable factors to be dependent on the same three characteristics. Therefore, the strong abnormal returns associated with momentum and idiosyncratic volatility are driven by time-varying betas related to these characteristics rather than by mispricing in alphas related to these characteristics.

Second, a standard feature of commodities markets is that they are somewhat segmented from stock and bond markets as well as from each other. Earlier empirical evidence reports low correlations of commodity returns with stock and bond returns and low correlations among commodity returns themselves (Gorton and Rouwenhorst, 2006; Erb and Harvey, 2006). This has changed dramatically in the past 20 years. The phenomenon has been referred to as the financialization of commodity markets (Tang and Xiong, 2012; Cheng and Xiong, 2014; Henderson et al., 2015; Brogaard et al., 2019). In the financialization process, there is a significant inflow of investment funds into commodity futures contracts. As a result, the price dynamics of commodity futures contract returns also change. In particular, commodity prices tend to be more synchronous. Commodity prices share a common boom and bust cycle. These new developments justify the use of latent factor models for commodity markets.

To examine whether this financialization process impact our results, we split the sample period into two sub-samples. The first sub-sample covers January 1981 to December 2002. The second sub-sample covers January 2003 to June 2022. Indeed, we find that the performance of IPCA latent factor models is much better in the second sub-sample. R^2's from the IPCA models are much larger in the second sub-sample than in the first sub-sample.

Then we move on to test for the significance of 20 individual commodity futures contract characteristics. The results are generally consistent with the evidence from the full sample, with a few exceptions. The sub-sample IPCA results indicate that in addition to momentum, expected shortfall, and idiosyncratic volatility, 12-month basis momentum has become significant in the first sub-sample. An indicator variable for 52-week high has become significant in the second sub-sample.

Third, when we compare the performance of standard models with observable risk factors with latent factor models, we find that the largest R^2 is 0.378 in a model that employs seven observable factors and seven commodity futures contract characteristics to instrument betas to these observable risk factors. In contrast, R^2's are 0.393 and 0.410, respectively, in the three-factor IPCA model with four and seven contract characteristics, but without any observable factors. Our evidence suggests that IPCA models dominate observable factor models in all model specifications.

Fourth, we find that some observable risk factors such as the momentum factor, the hedging pressure and speculative pressure factor, and the idiosyncratic volatility factor provide additional benefits in explaining commodity futures returns beyond the baseline IPCA model with the three latent factors. We also show that the momentum factor, the hedging pressure and speculative pressure factor, and the skewness factor are not fully spanned by the three latent factors when we implement the spanning test, as in Barillas and Shanken (2017) and Fama and French (2018).

Fifth, consistent with the above results, we further show that a three latent factor IPCA model together with seven observable factors can generate Sharpe ratios of 0.737 and 0.703, respectively, using four and seven futures characteristics to instrument betas of both observable and latent risk factors. These Sharpe ratios are higher than the Sharpe ratio of 0.668 from the Fama and French (2015) five-factor model plus the momentum factor for the U.S. equity market.

The rest of the paper proceeds as follows. Section 2 provides a brief review of the related literature. Section 3 introduces the IPCA procedure. Section 4 describes data sources and variable definitions. Section 5 reports summary statistics. Section 6 presents the main empirical results, including standard hedge portfolio returns sorted on commodity futures characteristics, the estimation of the baseline IPCA model, the significance of individual characteristics, the IPCA model with observable risk factors, spanning tests, and Sharpe ratios. Section 7 offers some additional discussion. Section 8 concludes the paper.

2. Literature Review

2.1 Studies on Commodity Futures Returns

There is growing evidence that long-short strategies based on exposure to commodity futures characteristics earn significant risk premiums. The literature has documented anomalous returns generated by sorting on commodity futures characteristics. Miffre (2016) provides a comprehensive review of various investment strategies in the commodity futures market. The widely-used sorting variables include (a) momentum (Miffre and Rallis, 2007; Fuertes et al., 2010, 2015; Narayan et al., 2014; Bianchi et al., 2015; Bakshi et al., 2019); (b) contrarian strategies (Asness et al., 2013; Bakshi et al., 2019); (c) 52-week high and low returns

(Bianchi et al., 2016); (d) volatility measured as the coefficient of variation or sum of squared daily returns (Gorton et al., 2013; Szymanowska et al., 2014); (e) idiosyncratic volatility (Fuertes et al., 2015); (f) skewness (Fernandez-Perez et al., 2018); (g) basis (Gorton et al., 2013; Yang, 2013; Szymanowska et al., 2014); (h) hedging pressure (De Roon et al., 2000; Basu and Miffre, 2013; Dewally et al., 2013); (i) basis momentum (Boons and Prado, 2019); (j) commodity market beta, USD beta, and inflation beta (Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006); (k) open interest (Hong and Yogo, 2012); and (l) liquidity (Szymanowska et al., 2014).

Bianchi et al. (2015) study trading strategies that jointly exploit momentum and contrarian strategies in commodity futures returns. The strongest abnormal returns come from a 12-month momentum together with an 18-month contrarian strategy. Fuertes et al. (2015) design a triple-screen trading strategy based on the three signals from momentum, term-structure (roll yields), and idiosyncratic volatility. Fernandez-Perez et al. (2019) compare the naïve equal-weight integration with six other more sophisticated style-integration approaches. They conclude that portfolios adopting the naïve equal-weight integration are unsurpassed by portfolios adopting other style-integration methods in terms of reward-to-risk profile. Fuertes and Zhao (2023) propose a Bayesian optimized style-integration strategy and show that this strategy achieves better Sharpe ratios and certainty equivalent returns.

Several papers examine the predictability of returns on commodity futures and test various versions of asset pricing models. These studies typically use macroeconomic or financial variables to test the predictability of either commodity market indices or individual commodity futures contracts (Bessembinder and Chan, 1992; Gargano and Timmermann, 2014; Ahmed and Tsvetanov, 2016; Daskalaki et al., 2014; Gao and Nardari, 2018; Koijen et al., 2018; and Baba Yara et al., 2019). Sakkas and Tessaromatis (2020) examine the performance of volatility timing strategies applied to a multi-factor commodity futures portfolio.

2.2 Firm Characteristics and Latent Factor Models

The asset pricing literature has identified many variables that can predict the cross-section of stock returns (Fama and French, 1993; Haugen and Baker, 1996; Subrahmanyam, 2010; Hou et al., 2015; Harvey et al., 2016). The literature has also investigated the source of these predictable patterns. Predictability can arise from either the ability to predict the cross-section of systematic risk (beta) or the ability to predict asset mispricing (alpha).

Daniel and Titman (1997) attempt to disentangle the alpha versus beta effect by double-sorting on lagged beta estimates and firm characteristics. Kim et al. (2021) argue that estimates of lagged beta are imprecise and stale. In addition, the double-sorting approach cannot handle two or more characteristics simultaneously. They extend the projected principal component analysis (PPCA) of Fan et al. (2016) to address the problems with the double-sorting approach. Latent factor betas are obtained from projected demeaned returns via standard principal component analysis (PCA). As a consequence, the obtained latent factors, as well as associated estimates of alphas and betas, depend on firm characteristics.

Kelly et al. (2019) propose an alternative latent factor model that also allows alpha and beta to be dependent on firm characteristics. Their instrumented principal component analysis (IPCA) considers alpha and beta to be linear functions of firm characteristics. The model allows the characteristics to change month by month, but the cross-sectional relationship between characteristics and either alpha or beta is constant over the full sample period. In contrast, the model of Kim et al. (2021) estimates beta using a rolling-window approach. As such, the cross-sectional relationship between either alpha or beta and characteristics varies over time. The Kim et al. (2021) model tends to perform better when the relationship between characteristics and either alpha or beta changes over time. For example, this happens when momentum itself is arbitraged away after its discovery (McLean and Pontiff, 2016).

2.3 Risk Factors

A number of earlier studies identify several commodity market risk factors that are priced. Szymanowska et al. (2014) examine the single basis factor. Yang (2013) considers a two-factor model with both the average commodity factor and the basis factor included. The average commodity factor is formed from an equal-weighted portfolio of all commodities. Boons and Prado's (2019) version of a two-factor model includes the average commodity factor and the basis momentum factor. Bakshi et al. (2019) mainly investigate a three-factor model that includes the average commodity factor, the basis (carry) factor, and the momentum factor.[3] They also include two additional factors: the value factor and the volatility factor. Boons and Prado (2019) test the performance of a four-factor model that includes the basis momentum

[3] Bakshi et al. (2019) construct a basis (carry) factor as the return on a portfolio that is long commodities that are the most backwardated (i.e., the lowest $\ln(y_t) < 0$) and short the ones that are the most in contango (i.e., the highest $\ln(y_t) > 0$), where $y_t = F_t^2 / F_t^1$ is the basis or slope of the futures curve and F_t^1 and F_t^2 are the first- and second-nearby futures contract prices.

factor in addition to the average commodity, basis, and momentum factors. Sakkas and Tessaromatis (2020) find evidence that a six-factor model contains all economically relevant pricing information. The six-factor model augments the four-factor model with two additional factors: the hedging pressure factor and the value factor.

3. The IPCA Model

3.1 Estimation of the IPCA Model

We present a brief overview of the IPCA estimation in this section. Kelly et al. (2019, 2020) provide a detailed analysis of the IPCA model.[4] This model supposes that there exist N commodity futures contracts for T periods. We specify the excess return on a commodity futures contract $r_{i,t+1}$ in the general IPCA framework as:

$$r_{i,t+1} = \alpha_{i,t} + \beta_{i,t} f_{t+1} + \varepsilon_{i,t+1}, \tag{1}$$

and

$$\alpha_{i,t} = z_{i,t}^{'}\Gamma_{\alpha} + v_{\alpha,i,t}, \quad \beta_{i,t} = z_{i,t}^{'}\Gamma_{\beta} + v_{\beta,i,t}, \tag{2}$$

where f_{t+1} is a $K\times 1$ vector of latent factors, $\alpha_{i,t}$ measures possible mispricing not captured by factors, and $\beta_{i,t}$ is a $1\times K$ vector of betas to latent factors. Both $\alpha_{i,t}$ and $\beta_{i,t}$ depend on observable commodity characteristics in the $L\times 1$ instrument vector $z_{i,t}$.

The $L\times K$ matrix Γ_{β} maps the potentially large number of characteristics to a small number of risk factor exposures. It allows us to condition the systematic risk estimates on lagged values of commodity futures characteristics. In doing so, the model can handle many characteristics simultaneously through the estimation of the matrix Γ_{β}. The residual $v_{\beta,i,t}$ captures any dynamic factor loading behavior which is orthogonal to characteristics. The $L\times 1$ vector Γ_{α} maps characteristics to any mispricing related to characteristics. The residual $v_{\alpha,i,t}$ captures any mispricing that is orthogonal to characteristics. By using a vector form and combining Equations (1) and (2), we have:

$$r_{t+1} = Z_t\Gamma_{\alpha} + Z_t\Gamma_{\beta} f_{t+1} + \varepsilon_{t+1}, \tag{3}$$

[4] The data and codes used in our study are available upon request.

where r_{t+1} is a $N\times 1$ vector of individual commodity futures returns at t, Z_t is a $N\times L$ matrix of commodity futures characteristics, and ε_{t+1} is a $N\times 1$ vector of composite errors, including error terms in both alphas and betas. We consider two alternative cases for the specification of Γ_α. The first is the restricted version in which characteristics do not proxy for alpha, i.e., $\Gamma_\alpha = 0$. The second is the unrestricted version in which characteristics do proxy for alpha, i.e., $\Gamma_\alpha \neq 0$.

Under the restricted specification $\Gamma_\alpha = 0$, the efficient estimators for Γ_β and f_{t+1} are obtained by minimizing the sum of squared compound errors:

$$\min_{\Gamma_\beta, f_{t+1}} \sum_{t=1}^{T-1} \left(r_{t+1} - Z_t \Gamma_\beta f_{t+1}\right)' \left(r_{t+1} - Z_t \Gamma_\beta f_{t+1}\right). \tag{4}$$

Kelly et al. (2019) show that the values of Γ_β and f_{t+1} that minimize Equation (4) satisfy the following two first-order conditions:

$$f_{t+1} = \left(\Gamma_\beta' Z_t' Z_t \Gamma_\beta\right)^{-1} \Gamma_\beta' Z_t' r_{t+1}, \forall t \tag{5}$$

and

$$vec(\Gamma_\beta') = \left(\sum_{t=1}^{T-1} Z_t' Z_t \otimes f_{t+1} f_{t+1}'\right)^{-1} \left(\sum_{t=1}^{T-1} \left[Z_t \otimes f_{t+1}'\right]' r_{t+1}\right). \tag{6}$$

Under the unrestricted specification $\Gamma_\alpha \neq 0$, we simply augment the factor specification to include a constant, $\tilde{f}_{t+1} = [1, f_{t+1}']'$, with the corresponding parameter matrix of $\Gamma = [\Gamma_\alpha, \Gamma_\beta]$. We then use the same procedure to estimate the parameters. Replacing f_{t+1} with $\tilde{f}_{t+1}$, we derive the similar first-order condition for Γ as in Equation (6):

$$vec(\Gamma') = \left(\sum_{t=1}^{T-1} Z_t' Z_t \otimes \tilde{f}_{t+1} \tilde{f}_{t+1}'\right)^{-1} \left(\sum_{t=1}^{T-1} \left[Z_t \otimes \tilde{f}_{t+1}'\right]' r_{t+1}\right). \tag{7}$$

The first order condition for f_{t+1} in Equation (5) changes slightly to:

$$f_{t+1} = \left(\Gamma_\beta' Z_t' Z_t \Gamma_\beta\right)^{-1} \Gamma_\beta' Z_t' (r_{t+1} - Z_t \Gamma_\alpha), \quad \forall t. \tag{8}$$

It is possible to nest observable factors such as commodity market-wide returns (*CMKT*), commodity market momentum (*CMOM12*), basis risk (*CBASIS*), hedging pressure and speculative pressure (*CHP*), basis momentum (*CBASM12*), idiosyncratic volatility (*CIVOL*), and skewness (*CSKEW*) into the IPCA framework. The encompassing model is:

$$r_{i,t+1} = \beta_{i,t} f_{t+1} + \delta_{i,t} g_{t+1} + \varepsilon_{i,t+1}, \tag{9}$$

where $\beta_{i,t}$ is specified as in Equation (1), $\delta_{i,t}$ is a $1 \times M$ vector of betas to observable factors, and g_{t+1} is a $M \times 1$ vector of observable factors which describes the portion of returns in addition to the IPCA latent factors f_{t+1}. Using a vector form, Equation (9) becomes:

$$r_{t+1} = Z_t \Gamma_\beta f_{t+1} + Z_t \Gamma_\delta g_{t+1} + \varepsilon_{t+1}. \tag{10}$$

Notice, in Equation (10), we impose the zero-alpha restriction $\Gamma_\alpha = 0$ to focus on exposure to systematic risks Γ_β and Γ_δ. Now we can specify $f_{t+1} = [f'_{t+1}, g'_{t+1}]'$ with the corresponding risk matrix $\Gamma = [\Gamma_\beta, \Gamma_\delta]$. The estimation for Γ is the same as in Equation (7). The estimation for f_{t+1} in Equation (8) now becomes:

$$f_{t+1} = \left(\Gamma'_\beta Z'_t Z_t \Gamma_\beta\right)^{-1} \Gamma'_\beta Z'_t (r_{t+1} - Z_t \Gamma_\delta g_{t+1}), \quad \forall t. \tag{11}$$

3.2 Hypothesis Tests in the IPCA Model

Equation (3) specifies the IPCA model with latent factors only. We carry out hypothesis tests after obtaining the estimates for the latent risk factors and key parameter matrix Γ_α and Γ_β. Our empirical analysis mainly focuses on three hypothesis tests. The first distinguishes between the unrestricted and restricted versions of the IPCA model, i.e. the zero alpha condition. The second tests for the significance of an individual characteristic, such as basis, while simultaneously controlling for all other characteristics. The third tests for whether each observed factor has incremental explanatory power. For example, we can test whether the basis momentum risk factor can also significantly contribute to the IPCA model when it is included.

The first hypothesis test aims to answer the question of whether commodity futures characteristics capture the difference in average returns that are unrelated to exposure to latent factors. In other words, the intercept in the latent factor model $Z_t\Gamma_\alpha$ is unrelated to Z_t. Specifically, we test the following null hypothesis:

$$H_0 : \Gamma_\alpha = 0_{L\times 1}$$

against the alternative hypothesis

$$H_A : \Gamma_\alpha \neq 0_{L\times 1}.$$

Using actual commodity futures return data, we construct a Wald-like statistic for the distance between alpha estimates from the unrestricted model and the restricted model ($0_{L\times 1}$) as the sum of squared elements in the Γ_α vector:

$$W_\alpha = \Gamma_\alpha' \Gamma_\alpha .$$

We follow Kelly et al. (2019) to construct the Wald statistic with bootstrapped data. A detailed discussion of each step in the bootstrapping process is provided in Appendix C. The basic idea is to sample residuals of characteristic-weighted portfolio returns from the unrestricted model. Then we impose the restriction to obtain predicted values of characteristic-weighted portfolio returns. The sum of the two components becomes the bootstrapped characteristic-weighted portfolio return, which we use to again estimate the unrestricted model to obtain bootstrapped estimates of Γ_α, Γ_α^{boot}. We repeat the bootstrapping process 1,000 times. The percentage of $W_\alpha^{boot} = \left(\Gamma_\alpha^{boot}\right)'\left(\Gamma_\alpha^{boot}\right)$ larger than $W_\alpha = \Gamma_\alpha' \Gamma_\alpha$ is the *p*-value for testing the null hypothesis. As argued in Kelly et al. (2019), the advantage of the bootstrap process is that we only resample residuals of characteristic-weighted portfolio returns from the unrestricted model. We do not need to resample contract level residuals.

The second test we implement is to evaluate the significance of an individual characteristic while simultaneously controlling for all other characteristics. We investigate whether a given characteristic significantly contributes to $\beta_{i,t}$. Consider the mapping matrix from characteristics to latent factors $\Gamma_\beta = [\gamma_{\beta,1}, ..., \gamma_{\beta,L}]'$, where $\gamma_{\beta,l}$ is a $K \times 1$ vector that maps the l^{th} element of characteristic vector $z_{i,t}$ to loadings on the K factors. The null and alternative hypotheses can be specified as follows:

$$H_0 : \Gamma_\beta = [\gamma_{\beta,1}, ..., \gamma_{\beta,l-1}, 0_{K\times 1}, \gamma_{\beta,l+1}, ..., \gamma_{\beta,L}]',$$

$$H_A : \Gamma_\beta = [\gamma_{\beta,1}, ..., \gamma_{\beta,l-1}, \gamma_{\beta,l,}, \gamma_{\beta,l+1}, ..., \gamma_{\beta,L}]'.$$

Under the null hypothesis, the entire l^{th} row of Γ_β is zero. The alternative hypothesis allows for a non-zero contribution from the l^{th} characteristic. If the null is rejected, the distance between the estimate of $\gamma_{\beta,l}$ under the alternative and the estimate of $\gamma_{\beta,l}$ imposed to be 0 under the null should be statistically large. The Wald-like statistic from using actual commodity futures return data is constructed as:

$$W_{\beta,l} = \gamma_{\beta,l}^{'} \gamma_{\beta,l} .$$

The inference of this test is to sample residuals of characteristic-weighted portfolio returns from the unrestricted model. Then we impose the restriction to obtain predicted values of characteristic-weighted portfolio returns. The sum of the two components becomes the bootstrapped characteristic-weighted portfolio return that we again use to estimate the unrestricted model to obtain the bootstrapped Wald-like statistic $W_{\beta,l}^{boot} = \left(\gamma_{\beta,l}^{boot}\right)^{'} \left(\gamma_{\beta,l}^{boot}\right)$.

The third test we implement is to examine whether a particular observable factor has incremental explanatory power after controlling for latent factors in the IPCA model and other observable factors. The hypotheses for this test are:

$$H_0 : \Gamma_\delta = [\gamma_{\delta,1}, ..., \gamma_{\delta,m-1}, 0_{L\times 1}, \gamma_{\delta,m+1}, ..., \gamma_{\delta,M}]',$$

$$H_A : \Gamma_\delta = [\gamma_{\delta,1}, ..., \gamma_{\delta,m-1}, \gamma_{\delta,m}, \gamma_{\delta,m+1}, ..., \gamma_{\delta,M}]'.$$

Similarly, the Wald-like statistic from using actual commodity futures return data is constructed as:

$$W_{\delta,m} = \gamma_{\delta,m}^{'}\gamma_{\delta,m}.$$

Using a similar bootstrapping procedure, we obtain the bootstrapped Wald-like statistic $W_{\delta,m}^{boot} = \left(\gamma_{\delta,m}^{boot}\right)^{'}\left(\gamma_{\delta,m}^{boot}\right)$. If $W_{\delta,m}$ is large relative to $W_{\delta,m}^{boot}$ for more than 950 of the 1,000 bootstraps, we can conclude that, at a significance level of 5%, the m^{th} observable factor of g_{t+1} has incremental explanatory power for individual commodity futures returns beyond the latent factors in the IPCA model and other observable factors

4. Data Sources, Commodity Futures Contracts, and Variable Definitions

4.1 Data Sources

Our main data source is Refinitive Eikon. We collect the S&P Goldman Sachs commodity total return indices (GSCI) on individual commodity futures contracts. The S&P GSCI is a widely used benchmark for investments in commodity futures. Earlier studies use GSCI individual commodity futures index data (Wang and Yu, 2004; Miffre and Rallis, 2007; Marshall et al., 2008; Bianchi et al., 2015, 2016). In addition, we use continuous commodity futures price series that are pre-constructed by the data vendor. This also follows the earlier literature.

When compiling continuous time-series futures returns, earlier studies use an immediate roll approach. All positions in the expiring futures contract are closed out on the same day as new positions are opened in nearby or distant contracts. S&P GSCI data compile individual commodities time-series prices by gradually rolling all futures positions over a pre-defined five-day period in each month. The approach is more practical for investors because rolling large positions on a single day may create an adverse price impact.[5] Gao and Nardari (2018) suggest that the GSCI Total Return Index measures a fully collateralized commodity futures investment. Its characteristics make it fully investable without any trading in physical commodities.[6]

The S&P GSCI futures data from Refinitive Eikon go back to 1974, but data are only available on a small number of futures contracts in the early years. We implement our empirical

[5] Please see Footnote 10 of Bianchi et al. (2016) for an example of an S&P GSCI gradual roll approach to construct the continuous time-series price data.

[6] Gao and Nardari (2018) also suggest that GSCI is by far the leading fully collateralized investable commodity index followed by exchange-traded products. It has the longest data history.

analysis starting in 1981 to guarantee that there are at least 10 commodities per month. Therefore, our sample of monthly GSCI indices covers the period of January 1981 to June 2022.

We also obtain daily settlement prices on first-nearby and second-nearby contracts from Refinitive Eikon. These daily data are used to construct the basis of commodity futures. The hedge and speculation positions are downloaded from the Commodity and Futures Trading Commission webpage. The trade weighted USD index returns and producer price index are from the Federal Reserve Bank at St. Louis's webpage.

Our dataset covers 34 commodity futures contracts that fall into the following five major categories: energy, grains and oilseeds, livestock, metals, and softs. The coverage is similar to that used by Hong and Yogo (2012), Gorton et al. (2013), Szymanowska et al. (2014), Bakshi et al. (2019), and Sakkas and Tessaromatis (2020). Some commodity futures contracts used in these earlier studies are not included in our sample because some data items are not available. For example, prices for the second-nearest futures contract to expiration are not available for propane, pork belly, rough rice, and milk and, therefore, the basis cannot be constructed. Similarly, the GSCI individual futures contract index for ethanol only starts in 2019 and is too short for empirical analysis.

We consider a total of 20 commodity futures characteristics. All characteristics are measured over a period prior to the prediction month *t*. For example, 12-month momentum return (*MOM12*) is measured as the cumulative return over prior months *t-12* to *t-1*. As commodity spot markets are known to be illiquid, we use the nearest to-maturity (first-nearby) futures price as the spot price, similar to Szymanowska et al. (2014) and most other studies on commodity futures. Daily basis is measured as the second-nearby futures contract daily price divided by the first-nearby futures contract daily price minus one. The detailed procedure for measuring these characteristics is provided in Appendix A.

Following Kelly et al. (2019), we transform these original measures of characteristics period by period into relative ranking in the cross-section then implement the IPCA estimation. Specifically, we first rank commodities on each characteristic, then rescale the rank by dividing it by the number of non-missing observations and subtracting 0.5. The procedure transforms the raw data on each characteristic into [-0.5, +0.5] intervals and focuses on their ordering rather than their magnitude. These standardized ranks should be insensitive to outliers.

We also construct a total of seven observable commodity market risk factors. These include the market-wide risk factor (*CMKT*), the 12-month momentum risk factor (*CMOM12*), the basis risk factor (*CBASIS*), the hedging pressure and speculative pressure risk factor (*CHP*), the basis momentum risk factor (*CBASM12*), the idiosyncratic volatility risk factor (*CIVOL*), and the skewness risk factor (*CSKEW*). The *CMKT* risk factor is measured as the monthly return on the GSCI overall commodity market index. For the *CMOM12*, *CBASIS*, *CBASM12*, *CIVOL*, and *CSKEW* risk factors, each is based on the rankings of commodity futures characteristics i.e., *MOM12*, *BASIS*, *BASM12*, *IVOL*, and *SKEWNESS*, respectively, in month *t-1* prior to the month *t* when we form equal weighted portfolio returns. The cutoff points for the high and low measures on commodity futures contract characteristics are top 30% and bottom 30% in each month *t-1*. The *CHP* risk factor is constructed from double-sorting on two variables *HPHE* and *HPSP*, respectively, where *HPHE* is hedging pressure and *HPSP* is speculative pressure. As in Fernandez-Perez et al. (2018), the long position consists of futures contracts with the highest *HPSP* (top 70%) and lowest *HPHE* (bottom 30%). The short position consists of futures contracts with the lowest *HPSP* (bottom 30%) and highest *HPHE* (top 70%).[7]

5. Summary Statistics

Table 1 reports summary statistics for our sample of 34 commodity futures contracts. The last three rows provide summary statistics for the commodity market index return, USD index return, and commodity produce price index (PPI) return. The complete time-series sample covers 498 months from January 1981 to June 2022. The commodity market index has an annual return of 3.8% with an annual standard deviation of 20.5%. The annualized Sharpe ratio is 0.071, similar to that reported in Gao and Nardari (2018). The three commodities that deliver the highest Sharpe ratios over the sample period are RBOB gasoline, soybean meal, and palladium, with corresponding Sharpe ratios of 0.327, 0.274, and 0.267, respectively.

Panels A1 and A2 of Table 2 provide summary statistics for 20 commodity futures characteristics. Panel A1 of Table 2 shows that the mean value of 12-month momentum measure (*MOM12*) is 0.030, or 3.0% per month. 12-month momentum is the cumulative return over prior months from *t-12* to *t-1*. The mean value of expected shortfall (*ES*) is -0.036, or -3.6% per month. Expected shortfall is a commonly used measure of the tail risk of financial

[7] The contrarian (i.e., value) strategy suggests that a portfolio of contracts with high measures of *CTR*36 or *CTR60* should have low returns, where *CTR36* and *CTR60* are 36- and 60-month cumulative returns from *t*-1 to *t*-36 and from *t*-1 to *t*-60, respectively. The long-short portfolio return should be negative. However, the results are opposite when we use *CTR60*. In addition, adding or deleting a few additional years of data surrounding our sample period from January 1981 to June 2022 does not generate significant long-short portfolio returns either. The evidence indicates the value effect is weak and, therefore, we do not include the value factor in the empirical analysis.

assets and is calculated as the average of the worst 5% returns. Here we measure the expected shortfall of individual futures contracts using the worst 5% of daily returns during prior months from *t-12* to *t-1*. Idiosyncratic volatility (*IVOL*) is measured as the standard deviation of the residuals from a regression of daily returns on commodity market index daily returns over prior months from *t-12* to *t-1*. The mean value of *IVOL* is 0.014 from our pooled time-series and cross-sectional sample.[8]

Panel A2 of Table 2 reports the pairwise correlations among 20 commodity futures contract characteristics. The 12-month momentum return (*MOM12*) is highly significantly correlated with 18-month contrarian return (*CTR18*), 36-month contrarian return (*CTR36*), and a measure of 52-week high based on relative price ratio (*R52WH*), with correlations of 0.82, 0.52, and 0.68, respectively. The correlation between idiosyncratic volatility (*IVOL*) and maximum daily return (*MAX*) is positive and highly significant at 0.73, while the correlation between idiosyncratic volatility (*IVOL*) and expected shortfall (*ES*) is negative and highly significant at -0.70. Hedging pressure (*HPHE*) has a negative and highly significant correlation of -0.75 with speculative pressure (*HPSP*). There is a negative correlation of -0.32 between basis (*BASIS*) and 12-month basis momentum (*BASM12*). The correlation between trading volume and the Amihud liquidity (Amihud, 2002) measure is significantly negative at -0.03.

Panel B1 of Table 2 provides summary statistics for the seven observable risk factors *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*. The mean values of these risk factors are 3.8%, 10.7%, 6.2%, 7.8%, 5.4%, 9.3%, and 6.8% per annum, respectively. Long-short portfolio returns sorted on *BASIS*, *IVOL*, and *SKEWNESS* are negative. For these three characteristics, we add a negative sign to form the *CBASIS*, *CIVOL*, and *CSKEW* risk factors. As a result, all seven risk factors have positive mean returns.

Panel B2 of Table 2 tabulates the pairwise correlations between these seven risk factors. In general, the correlations are low. The highest correlation is 0.47 between *CBASIS* and *CBASM12*.

6. Empirical Results

6.1 One-way Sorted Commodity Futures Hedge Portfolio Returns

[8] We compute alternative measures of tail risk such as value-at-risk of daily returns on the GSCI individual commodity futures contract index at the 5% level. We also compute the summation of squared daily returns on the GSCI individual commodity futures contract index. But the IPCA model estimation using these two characteristics fails to converge.

We begin our empirical analysis with standard one-way sorted long-short portfolio returns. For each of the 20 commodity futures characteristics, Table 3 reports the mean returns on the long position, the mean returns on the short position, the mean returns on the hedge portfolios (long – short), *t*-statistics, and number of monthly observations. Among the 20 characteristics, we find 12 generate significant long-short portfolio returns at the 5% level: *MOM3*, *MOM12*, *R25WH*, *IVOL*, *SKEWNESS*, *MAX*, *BASIS*, *HPHE*, *HPSP*, *VOLM*, *OPNI*, and *ALIQ*. All 12 characteristics have expected signs based on the earlier literature.

The magnitude of monthly returns from these 12 hedge portfolios is large, ranging from -5.60% per annum to 11.40% per annum with corresponding *t*-statistics of -2.05 to 4.28, respectively. The 12-month momentum strategy (*MOM12*) generates an annual return of 10.70% with a *t*-statistic of 3.75. The strong momentum effect is consistent with those reported in earlier studies (Miffre and Rallis, 2007; Fuertes et al., 2010, 2015; Asness et al. 2013; Bianchi et al., 2015, 2016; and Bakshi et al., 2019).

McLean and Pontiff (2016) study post-publication return predictability of 97 variables shown to predict cross-sectional stock returns in the U.S. equity market.[9] They report an average predictor's long-short return shrinks 58% post-publication with a lower bound on the publication effect of about 32%. They reject both the hypothesis that return predictability disappears entirely and the hypothesis that post-publication return predictability does not change. We explore the same issue for commodity futures returns. For example, our results in Table 3 indicate that the long-short portfolio sorted on expected shortfall (ES) yields an annual return of 3.70% with a *t*-statistic of 1.36. But if we shorten our sample period from January 1981 to December 1998, the long-short portfolio generates an annual return of 8.90% with a highly significant *t*-statistic of 1.95.

The characteristic that deserves more discussion is the volatility measure. Some earlier studies use different measures of total volatility (Szymanowska et al., 2014) or explore the contemporaneous relation between total volatility and hedge portfolio returns (Gorton et al., 2013). We study the effect of idiosyncratic volatility measured at *t-1*, or one-month lag.

Our results in Table 3 show that trading futures contracts based on idiosyncratic volatility generates an annual return of -9.30% with a *t*-statistic of -3.37. Commodity futures with higher idiosyncratic volatility have a lower return than commodity futures with a lower

[9] McLean and Pontiff (2016) compare the return predictability of firm characteristics over three distinct periods: (i) the original study's sample; (ii) after the original sample but before publication; and (iii) post-publication. They attempt to differentiate between alternative explanations for the return predictability such as statistical biases, rational pricing, and mispricing.

idiosyncratic volatility. The results remain robust when we use a longer sample from January 1973 to December 2022, for a total of 594 months. The long-short portfolio sorted on *IVOL* generates an annual return of -8.26% with a *t*-statistic of -2.92. Our results are also consistent with those from Fuertes et al. (2015).

In an earlier study, Gorton et al. (2013) report a positive long-short hedge portfolio return sorting on contemporaneous volatility measured in month *t*.[10] Specifically, they demean volatility by the time-series mean of volatility, where volatility is calculated as the square root of the average squared daily excess returns of the month over which the excess return is calculated. Szymanowska et al. (2014) study the relationship between commodity futures returns on two trading strategies (i.e., short-roll returns, and excess holding returns) and volatility measured as a coefficient of variation (i.e., standard deviation divided by mean).[11] They report a positive relationship between volatility and short-roll returns and a negative relationship between volatility and excess holding returns.[12]

We also calculate the sum of daily squared returns from month *t-12* to month *t-1*. This is a measure of total volatility. The hedge portfolio sorted on total volatility yields an annual return of -4.88% with a *t*-statistic of -1.86. If we use a longer sample from January 1973 to June 2022, for a total of 594 months, the hedge portfolio sorted on total volatility yields an annual return of -6.47% with a *t*-statistic of -2.29. The abnormal return from sorting on total volatility is in the same direction as the abnormal return from sorting on idiosyncratic volatility.[13]

Our conclusion regarding idiosyncratic volatility remains valid if we use a sample period from January 1973 to December 2010 which is close to the sample period of Gorton et al. (2013), a sample period from March 1986 to December 2010 that matches the sample period of Szymanowska et al. (2014), or a sample period from January 1979 to August 2011 that matches the sample period of Fuertes et al. (2015).

6.2 The Basic IPCA Model

---

[10] See Panel A in Table 9 of Gorton et al. (2013).

[11] A short-roll strategy invests in one-period futures contracts for *n* consecutive periods, that is, rolling them over each period. A holding strategy buys an *n*-period futures contract at time *t* and holds it until the maturity date *t* + *n*. The difference between holding period return and short-roll return is the excess holding return.

[12] See Panel C in Table 3 of Szymanowska et al. (2014).

[13] The correlation between total volatility and idiosyncratic volatility is high at 0.755 and 0.753, respectively, during the January 1973 to June 2022 and January 1981 to June 2022 periods.

The first item we need to determine is the number of latent factors in the IPCA model. Then we test whether we can restrict the intercept Γ_α in the IPCA model to be zero, in which case only betas to latent factors depend on commodity futures characteristics. In the unrestricted model $\Gamma_\alpha \neq 0$, both alphas and betas to latent factors depend on futures characteristics. Therefore, futures characteristics affect both mispricing and risk loadings and consequently affect realized returns. To test the null hypothesis of $\Gamma_\alpha = 0$ versus the alternative hypothesis $\Gamma_\alpha \neq 0$, we construct *p*-values by counting the percentage of the 1,000 bootstrapped Wald statistics $W_\alpha^{boot} = \left(\Gamma_\alpha^{boot}\right)' \left(\Gamma_\alpha^{boot}\right)$ that are larger than the Wald statistics $W_\alpha = \Gamma_\alpha' \Gamma_\alpha$ calculated using the actual data.

We estimate the IPCA model with different numbers of latent factors, i.e., $K = 1$, 2, and 3. There is a constraint in the IPCA model on how many characteristics can be included. We cannot include too many characteristics in the model because the number of contracts is small, i.e., fewer than 10, in the early years of our January 1981 to June 2022 sample period. Therefore, in the baseline model, we only include four characteristics (L=4), i.e., Z = (*MOM12*, *CTR36*, *ES*, *IVOL*). We also consider seven characteristics (L=7) in the IPCA model, i.e., Z = (*MOM12*, *CTR36*, *ES*, *IVOL*, *BASIS*, *BASM12*, *ALIQ*). Our results are robust to alternative choices of four or seven characteristics in the baseline model.

Table 4 summarizes our results for the baseline IPCA model. Panel A reports R^2's from restricted and unrestricted models for $K = 1$, 2, and 3 with four characteristics (L=4). The unrestricted models yield R^2's of 0.257, 0.340, and 0.394 for $K = 1$, 2, and 3, higher than R^2's from the restricted models. The Wald test rejects the null that $\Gamma_\alpha = 0$ with a *p*-value of 0.005 and 0.027 for $K = 1$ and 2. The Wald test fails to reject the null that $\Gamma_\alpha = 0$ with *p*-values of 0.806 for $K = 3$. Panel B of Table 4 reports the results for the IPCA model with seven characteristics (L=7). More characteristics do provide more information for explaining commodity futures returns with higher R^2's across different model specifications. The Wald test also fails to reject the null hypothesis that $\Gamma_\alpha = 0$ at the 5% level when $K = 3$. The corresponding *p*-value is 0.647.

Overall, the results in Table 4 indicate that with three latent risk factors, i.e., $K = 3$, IPCA essentially explains a significant portion of the variation in commodity futures returns (e.g., R^2 is 39.3% in a three-latent factor model with four characteristics). For IPCA models with an intercept, the incremental R^2 from adding the second latent factor is 8.5%. The incremental R^2 from adding the third latent factor is 5.4%. Our unreported results indicate that the incremental

R^2 from adding the fourth latent factor is 4.3%. The incremental R^2 from adding the fifth latent factor is 3.4%.

The null hypothesis that $\Gamma_\alpha = 0$ cannot be rejected at the 5% level when $K = 3$. Therefore, the impact of characteristics is negligible in the alphas of the latent factor models. For our subsequent analysis, we focus on the IPCA model with three latent factors and restrict alpha to zero.

6.3 The Significance of Individual Commodity Futures Characteristics

In this section, we test for the significance of individual characteristics. We have a total of 20 commodity futures characteristics. As mentioned earlier, we cannot include too many characteristics in the model because the number of contracts is small in the early years of our sample period. Therefore, the baseline model only contains four characteristics. We then add each of the remaining 16 characteristics, one at a time, and test for its significance. In the augmented model, we first include seven characteristics and then add each of the remaining 13 characteristics, one at a time, and test for its significance. We also experiment with alternative choices of characteristics in the baseline and the augmented model. Our conclusions regarding the significance of each of the 20 characteristics are robust. The significance of each characteristic is tested via the bootstrap method described in Kelly et al. (2019) and Appendix C. In both the baseline and the augmented model, we choose the restricted IPCA model with $\Gamma_\alpha = 0$ and the three latent factors, i.e. $K = 3$, based on the results from Table 4.

The first row in Panel A of Table 5 presents the bootstrapped *p*-value for each individual characteristic in the baseline IPCA model with four characteristics (L=4). Following the procedure outlined in Section 3.2 and Appendix C, *p*-values are constructed by counting how many of the 1,000 bootstrapped Wald statistics $W_{\beta,l}^{boot} = \left(\gamma_{\beta,l}^{boot}\right)' \left(\gamma_{\beta,l}^{boot}\right)$ are larger than the Wald statistics $W_{\beta,l} = \gamma_{\beta,l}' \gamma_{\beta,l}$ calculated using the actual data. The results indicate that three of the four characteristics are consistently significant: *MOM12*, *ES*, and *IVOL*. The corresponding *p*-values are 0.000, 0.003, and 0.000, respectively. The exception is *CTR36*, with a bootstrapped *p*-value of 0.742. Then we add each of the 16 remaining characteristics, one at a time, into the IPCA model. The last column in Panel A shows that 15 out of these 16 characteristics are not significant. The only exception is *MOM3*. This is not surprising as it has a highly significant correlation of 0.52 with *MOM12*.

Next, we test the augmented IPCA model with seven characteristics (L=7) and then add additional characteristics, one at a time, in subsequent tests. The first row in Panel B indicates a similar pattern. Those three significant characteristics in the baseline model, i.e., *MOM12*, *ES*, and *IVOL*, retain their significance in the augmented model. The corresponding *p*-values are 0.000, 0.031, and 0.003, respectively. The corresponding *p*-values for *CTR36*, *BASIS*, *BASM12*, and *ALIQ* are 0.796, 0.555, 0.870, and 0.437, respectively. Then we add each of the 13 remaining characteristics, one at a time, into the IPCA model. The last column in Panel B suggests that none of these 13 characteristics is significant, including *MOM3*.

These results in Table 5 show that among the 20 characteristics, only a few make a substantial contribution to explaining the betas in the latent factor model. Controlling for *MOM12*, *ES*, and *IVOL*, none of the remaining characteristics deliver even a significantly marginal contribution. This result is similar to what Kelly et al. (2019) find for the U.S. equity market. They report that of the 36 characteristics, only ten are significant at the 1% level. Among the ten significant characteristics, only two stand out for the magnitude of their marginal contribution to the model.

6.4 The Significance of Observable Risk Factors

Now we turn to estimate the IPCA model, including seven observable factors in the commodity futures market. We test for the significance of these observable factors. Just like the latent factors, the betas of these observable factors are also instrumented with futures characteristics, i.e., the betas of these observable risk factors depend on characteristics. We consider the market (*CMKT*), momentum (*CMOM12*), basis (*CBASIS*), hedging pressure and speculative pressure (*CHP*), basis momentum (*CBASM12*), idiosyncratic volatility risk (*CIVOL*), and skewness (*CSKEW*) factors in the commodity futures market. These popular commodity market risk factors have been used in a number of earlier studies (Yang 2013; Szymanowska et al., 2014; Bakshi et al., 2019; Boons and Prado, 2019; Sakkas and Tessaromatis, 2020).

Table 6 summarizes R^2's from the alternative model specifications. The sample covers the period from January 1987 to June 2022, for a total of 426 months. The sample period is shorter than our full sample period from January 1981 to June 2022. The reason is that the earliest CFTC long and short hedger and speculator positions are only available after January 1986. The construction of the hedging pressure and speculative pressure factor (*CHP*) requires the availability of long and short positions from hedgers and speculators. The number of latent factors ranges from one to three. We also restrict the intercept to be zero ($\Gamma_\alpha = 0$). Panel A employs four commodity futures characteristics: Z=(*MOM12*, *CTR36*, *ES*, *IVOL*). Panel B employs seven commodity futures contract characteristics: Z=(*MOM12*, *CTR36*, *ES*, *IVOL*, *BASM*, *BASM12*, *ALIQ*).

We first report the estimation results using four characteristics in the IPCA model. The first part of Panel A of Table 6 shows that the inclusion of observable risk factors strengthens the explanatory power. With a single latent factor, IPCA alone can explain 0.254 of total variation in commodity futures returns. The inclusion of seven observable risk factors raises R^2 to 0.390, making a marginal contribution of 0.136 (0.390 – 0.254) in R^2. As the number of latent factors increases, the marginal contribution from observable factors shrinks. With two latent factors, the increase in R^2 is 0.079. With three latent factors, the increase in R^2 is 0.045. These results show that observable risk factors do provide some benefits in explaining commodity futures returns, yet their explanatory power is much weaker than that from the IPCA latent factors. More importantly, as the number of latent factors increases, the additional explanatory power from observable factors decreases. Latent factors from the IPCA model extract useful information explaining individual futures returns more efficiently than observable risk factors.

The second part of Panel A of Table 6 tests for the significance of each of these seven observable factors. As shown in Section 3.2 and Appendix C, the *p*-value is constructed comparing the 1,000 bootstrapped Wald statistics $W_{\delta,m}^{boot} = \left(\gamma_{\delta,m}^{boot}\right)' \left(\gamma_{\delta,m}^{boot}\right)$ with the Wald statistics $W_{\delta,m} = \gamma_{\delta,m}' \gamma_{\delta,m}$ obtained from the actual data. The empirical evidence indicates that with one latent factor we can easily reject the null hypothesis that the individual observable factor does not provide additional explanation for six out of seven observable factors. With two latent factors, we can reject the null for five out of seven observable factors. With three latent factors, we can reject the null for three out of seven observable factors.

Panel B of Table 6 repeats the procedure including seven characteristics in the IPCA model. Overall, R^2's increase across all model specifications with seven characteristics relative to four characteristics, including models with latent factors only and models with both latent factors and observable factors. With one latent factor, we can reject the null for five out of seven observable factors. With two latent factors, we can reject the null for six out of seven observable factors. With three latent factors, we can reject the null for five out of seven observable factors.

There are some significant differences between the U.S. equity market and commodity futures markets. Kelly et al. (2019) conclude that observable factors start to become redundant with more IPCA factors. At K = 5, none of the Fama and French (2015) factors plus the momentum factor are statistically significant at the 1% level after controlling for IPCA factors. For the commodity futures market, we find that more observable factors are significant even controlling for latent factors. What contributes to these differences requires more exploration.

6.5 Significance of Individual Characteristics When Observable Factors Are Also in the IPCA Model

In Section 6.3, we test for the significance of individual characteristics and find that only three characteristics, i.e., *MOM12*, *ES*, and *IVOL*, are significant instruments for the betas of the latent factors in the IPCA model. The model in Section 6.3 excludes observable factors in implementing the tests. Once again, we test the significance of these three characteristics in the IPCA model. However, the difference is that now we also include the seven observable factors, i.e., *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASIS*, *CIVOL*, and *CSKEW*, in the IPCA model. In addition, the betas of these observable factors are also instrumented by *MOM12*, *ES*, and *IVOL* in the same way as the betas of the three latent factors are instrumented.

The bootstrapping procedure for obtaining *p*-values is similar to that used in Section 6.3, with some minor modifications. Essentially, we obtain the residuals of the characteristic-based portfolio returns under the alternative hypothesis. Then we construct the predicted value of the characteristic-based portfolio returns under the null hypothesis by imposing the restriction that the entire l^{th} row of Γ_β or $\gamma_{\beta,l}$ is zero. These two components add up to the bootstrapped returns. We estimate the IPCA model with observable factors again using these bootstrapped returns. Notice here the betas of observable factors still depend on the l^{th} characteristic, i.e., the entire l^{th} row of Γ_δ is not restricted to be zero.

Table 7 summarizes the bootstrapped *p*-values. We only report the results for the case of three latent factors, K=3, in the IPCA specification. There are a total of eight alternative combinations of observable factors in Models 1 to 8. In Model 1, for example, *CMKT* is the only observable factor included in the IPCA estimation. But the beta of *CMKT* is also instrumented by *MOM12*, *ES*, and *IVOL*. In this case, all three characteristics that are used to instrument betas of the three latent factors remain highly significant. Models 2 to 5 include 2, 3, 4, and 5 observable factors, respectively. Now *MOM12* becomes insignificant. But *ES* and *IVOL* remain highly significant. Notice Models 2 to 5 include the observable momentum factor

CMOM12. This addition significantly reduces the impact of characteristic *MOM12* in the betas of latent factors. When observable factor *CMOM12* is excluded in Model 6, the characteristic *MOM12* again becomes highly significant. Finally, Models 7 and 8 include six and seven observable factors, respectively. The last two rows in Table 7 indicate that all three characteristics remain highly statistically significant.

6.6 Financialization of Commodities Futures Market and Sub-Sample IPCA Estimation

The commodity futures market has experienced some significant changes over the past 20 years as more alternative investment managers have taken positions in commodity futures markets as a proxy for the cash market. This process is often referred to as the financialization of commodity markets (Tang and Xiong, 2012; Cheng and Xiong, 2014; Henderson et al., 2015; Brogaard et al., 2019). There exist a number of features to the process. First, there is a significant inflow of investment funds into commodity futures contracts. Portfolio managers have started to treat commodity futures as a standard separate asset class such as stocks and bonds. Second, commodity futures price dynamics have changed substantially. The correlations of commodity prices with prices in other asset classes such U.S. and emerging market equity rose significantly after 2003. In earlier years, commodity prices had little co-movement with stocks (Gorton and Rowenhorst, 2006) or each other (Erb and Harvey, 2006). Third, following the significant inflow of investment capital from commodity index traders, both gross and net positions in futures markets grew dramatically after 2003.

To examine whether this financialization process impacts our results, we split the sample period into two sub-samples. The first sub-sample covers January 1981 to December 2002. The second sub-sample covers January 2003 to June 2022. The first sub-sample contains 264 months. The second sub-sample contains 234 months. Panel A in Table 8 shows that the performance of the IPCA model is much better in the second sub-sample. R^2's from IPCA models with K = 1, 2, and 3 latent factors are much larger in the second sub-sample than in the first sub-sample. This is consistent with the financialization of commodity futures markets in the second sub-sample when individual commodity futures contract returns tend to be more synchronous.

In addition, Panels B and C in Table 8 test for the significance of individual futures contract characteristics. Both panels only include four characteristics in the model. The first three characteristics are *MOM12*, *ES*, and *IVOL*. The fourth characteristic is one of the remaining 20 - 3 = 17 characteristics. The results are similar to the full sample with a few exceptions. The results indicate that in addition to *MOM12*, *ES*, and *IVOL*, *MOM3* and *BASM12* are significant in the first sub-sample. *MOM3* and *R52WH* are significant in the second sub-sample.

6.7 Comparing Alternative Asset Pricing Models

Previous sections focus on the estimation and significance tests of the IPCA model with and without observable factors. In this section, we compare the performance of these models in explaining cross-sectional variation of commodity futures returns. For that purpose, we first specify the IPCA models with betas being instrumented by four commodity futures characteristics (L=4) with and without observable risk factors. We then allow the IPCA models with betas being instrumented by seven commodity futures contract characteristics (L=7) with and without observable risk factors.

Panel A of Table 9 summarizes R^2's from the two sets of IPCA models. We report the results for the number of latent factors to be one, two or three, i.e., K = 1, 2, and 3, but our discussion focuses on the case when K = 3, (i.e., the last column in Panel A). When there are four characteristics used as instruments (L=4), the IPCA model generates an R^2 of 0.393 when no observable factors are involved. IPCA models instrumented with four characteristics and the inclusion of three, four, five, six, and seven observable factors generate R^2's of 0.425, 0.427, 0.427, 0.438, and 0.438, respectively. When there are seven characteristics used as instruments (L=7), IPCA models generate an R^2 of 0.410 when no observable factors are involved. IPCA models instrumented with seven characteristics and the inclusion of three, four, five, six, and seven observable factors generate R^2's of 0.466, 0.469, 0.485, 0.491, and 0.491, respectively. These results indicate that IPCA models with observable factors outperform IPCA models without observable factors.

Panel B of Table 9 summarizes R^2's from three sets of observable factor models. We first consider the first set of observable factor models when betas are constant. The three, four, five, six, and seven-factor models generate R^2's of 0.143, 0.148, 0.149, 0.169, and 0.169, respectively. In the second set of observable risk factor models, we allow betas of these observable factors to be dependent on commodity futures characteristics. When there are four characteristics (L=4) in the model, the three, four, five, six, and seven-factor models generate

R^2's of 0.260, 0.268, 0.270, 0.322, and 0.325, respectively. The improvement over constant beta models is obvious.[14]

Now we allow seven characteristics (L=7) to instrument the betas in observable factor models. We require all seven characteristics be available. In the third set of observable factor models when betas are also constant, the three, four, five, six, and seven-factor models generate R^2's of 0.150, 0.154, 0.156, 0.174, and 0.175, respectively. In the fourth set of observable risk factor models when betas vary with contract characteristics, the three, four, five, six, and seven-factor models generate R^2's of 0.299, 0.307, 0.330, 0.376, and 0.378, respectively. The improvement over constant beta models is, again, obvious.

Overall, the largest R^2 from Panel B is 0.378 in the seven-factor model with L=7. R^2's from Panel A are 0.393 and 0.410, respectively, in the IPCA model with four and seven contract characteristics, but without any observable factors. Clearly, IPCA models dominate observable factor models in all specifications. These results confirm the overwhelming advantage IPCA models plus observable factors have relative to models with observable risk factors only.

6.8 Spanning Tests and Sharpe Ratios

6.8.1 Spanning Tests

Barillas and Shanken (2017) and Fama and French (2018) use spanning regression analysis to assess the benefits from adding a factor to an existing factor model. The methodology regresses a candidate factor on the model's other factors. A nonzero intercept indicates the candidate factor is not spanned by other factors and makes a marginal contribution to the existing model. We explore whether latent factors are spanned by observable factors and vice versa. Specifically, we first regress each of the three latent factors *FAC1*, *FAC2*, and *FAC3* from the IPCA model on seven observable factors, i.e., *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*. By doing so, we test whether each latent factor can be spanned by the seven observable factors. Next, we re-run the spanning regression by switching the independent and dependent variables. We set the three latent factors as independent variables and regress each observable factor *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW* on the three latent factors *FAC1*, *FAC2*, and *FAC3*. This allows us to test whether each observable factor can be spanned by the three latent factors. Finally, we compare

[14] We test whether each of the 20 commodity futures characteristics affects the betas of the seven observable factors. The null hypothesis can be firmly rejected for each of the four characteristics (*MOM12*, *CTR36*, *ES*, and *IVOL*) at the 5% significance level. The null hypothesis can also be firmly rejected at the 5% significance level for 14 out of 16 remaining characteristics.

the annualized Sharpe ratios for various portfolios formed from latent factors, observable factors, and a combination of latent and observable factors, respectively.

Panels A1 and A2 of Table 10 show that the intercepts in the spanning regressions for each of the three latent factors are not significant at the 5% level. All three latent factors are actually well-spanned by the seven observable factors.

Panel A1 also indicates that the first latent factor (*FAC1*) is mostly related to market-wide movement and idiosyncratic volatility in commodity futures contracts. The estimated coefficient on *CMKT* is 1.355 with a *t*-statistic of 11.52. The estimated coefficient on *CIVOL* is 1.497 with a *t*-statistic of 12.07. The second latent factor (*FAC2*) is mostly related to the momentum factor. The estimated coefficient (*t*-statistic) on *CMOM12* is 1.216 (18.75). These two strong patterns remain robust in Panel A2 when we add more characteristics as instruments.

In Panels B1 and B2 of Table 10, we repeat the spanning regression tests by switching the independent and dependent variables. From Panel B1, we find that *CMKT*, *CHP* and *CSKEW* are not fully spanned by the three latent factors from the model when we use four characteristics to instrument betas of the latent factors. From Panel B2, we find that *CMMOM12*, *CBASIS*, *CHP*, and *CSKEW* are not fully spanned by the three latent factors from the model when we use seven characteristics to instrument betas of latent factors.

In addition, the empirical evidence from Panels B1 and B2 also confirms that *FAC1* captures movement in the market risk factor (*CMKT*) and idiosyncratic volatility factor (*CIVOL*). At the same time, *FAC2* captures movement in *CMOM12*. From Panels B1 and B2, the *t*-statistics associated with *FAC1* are around 15.00 when the dependent variable is *CMKT*. The *t*-statistics associated with *FAC1* exceed 11.00 when the dependent variable is *CIVOL*. The *t*-statistics associated with *FAC2* exceed 16.00 when the dependent variable is *CMOM12*.

6.8.2 Sharpe Ratios

The Sharpe ratio provides a summary statistic for the trade-off between risk and return. This section compares Sharpe ratios from investment opportunities constructed using varying combinations of latent factors, observable factors, and both latent and observable factors in the commodity futures market. In addition, we provide Sharpe ratios from various combinations of risk factors from the Fama and French (1993, 2015) model. This gives us some sense of the difference in risk-return trade-offs from the U.S. equity market and the commodity futures market.

From Panel C of Table 10, we find that the three latent factors with four characteristics generate an annualized Sharpe ratio of 0.565, while the three latent factors with seven characteristics produce an annualized Sharpe ratio of 0.464. Moving to observable factors, three observable factors (*CMKT*, *CMOM12*, and *CBASIS*) generate an annualized Sharpe ratio of 0.439 while four observable factors (*CMKT*, *CMOM12*, *CBASIS*, and *CHP*) generate an annualized Sharpe ratio of 0.482. When all seven observable risk factors are employed, the Sharpe ratio reaches a value of 0. 649. In general, we find that a commodity portfolio based on the IPCA based latent factors generates a risk return trade-off slightly below that of a commodity portfolio based on observable factors (0.565 versus 0.649). Combining the three latent factors with seven observable factors generates an annualized Sharpe ratio of 0.737 and 0.703, respectively, when we use four and seven characteristics as instruments for betas to latent factors.

Finally, Panel C summarizes Sharpe ratios using various observable risk factors in the U.S. equity market. The Fama and French (1993) three factors, *EXMRET*, *SMB*, and *HML*, yields an annualized Sharpe ratio of 0.379. The Fama and French (2015) five factors, *EXMRET*, *SMB*, *HML*, *RMW*, and *CMA*, generates an annualized Shape ratio of 0.574. If we add the momentum factor (*MOM*) to the Fama and French (2015) five factors, the annualized Sharpe ratio becomes 0.668. Therefore, with latent factors and observable factors, commodity futures can offer a better risk-return trade-off than that from the U.S. equity market.

To provide a visual effect of risk and return trade-off from latent factors, observable factors, and U.S. equity market factors, Figure 1 plots the efficient frontiers constructed from the following combinations of assets: (i) *CMKT*, *CMOM12*, and *CBASIS*; (ii) *CMKT*, *CMOM12*, *CBASIS*, and *CHP* ; (iii) *CMKT*, *CMOM12*, *CBASIS*, *CHP*, and *CBASM12*; (iv) *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, and *CIVOL*; (v) *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*; (vi) *FAC1*, *FAC2*, and *FAC3* (L=4); (vii) *FAC1*, *FAC2*, and *FAC3* (L=7); and (viii) FF5 (*EXMRET*, *SMB*, *HML*, *RMW*, *CMA*) plus *MOM*.

Two strong patterns appear in Figure 1. First, the volatility of the efficient frontier from the three latent factors (*FAC1*, *FAC2*, and *FAC3*) is higher than the volatility of the efficient frontier from seven observable factors (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*), as with the mean returns. Second, commodity futures market contracts exhibit a much higher volatility than the U.S. equity market, as with the mean returns. However, the Sharpe ratio from the U.S. equity market is 0.668, slightly lower than ratios of 0.737 and 0.703

from combining the three latent factors and seven observable factors, respectively, in the commodity futures market.

7. Additional Discussion

7.1 Using Contract Characteristics to Predict Both Returns and Betas

There is general agreement in the existing asset pricing literature that asset characteristics can predict future asset returns. The controversial issue is whether this relationship is driven by *asset characteristics per se* or driven by *betas that are related to these characteristics*. In Panel A of Appendix D, we run panel regressions of commodity futures contract monthly returns on contract characteristics using pooled time-series and cross-sectional data. The returns are measured in month *t* while characteristics are measured from month *t*-12 to month *t*-1. The results indicate that 11 out of 20 contract characteristics are significant. The evidence is consistent with that from Table 3 using a portfolio approach.

In Panel B of Appendix D, we run the same panel regressions, but the dependent variable is the estimated market beta using only 12 observations from month *t*+1 to *t*+12. This follows Kelly et al. (2023). The purpose is to examine whether the same set of 20 contract characteristics can also predict market betas of futures contract returns. Market betas are estimated with respect to returns on the GSCI commodity market index. Now 9 out of 20 contract characteristics are highly significant. In particular, the estimated coefficients on *MOM12*, *ES*, and *IVOL* are stable and highly significant in different model specifications. *MOM12* is positively related to market beta. *ES* and *IVOL* are negatively related to market beta. R^2's from the panel regressions are also high, ranging from 0.234 to 0.299. Kelly et al. (2023) report an R^2 of 0.166 for a similar panel regression in the bond market.

Our evidence from Appendix D indicates that contract characteristics can predict both risk and returns of commodity futures contracts. However, the evidence in Appendix D cannot tell us how much of the return predictability is caused by the fact that contract characteristics can also predict market betas. The IPCA approach resolves this issue. It allows data to identify the most important latent factors and test whether these characteristics are, in fact, related to alphas or betas of latent factors.

In the IPCA estimation, we consider a total of 20 commodity futures contract characteristics. We find that only three are significantly related to betas. However, our evidence should be interpreted with care. The test for the significance of contract characteristics is a joint test for the model specification and the null hypothesis that the contract characteristic is not related to betas of latent factors. In the IPCA framework, the relation between beta and contract characteristics is linear. In addition, the linear relation is fixed for the entire sample period or estimation period. These restrictions are strong because existing evidence shows that the predictive power of characteristics changes over time (McLean and Pontiff, 2016; Kim et al., 2021). The consistency of IPCA estimates requires a large time-series sample. A rolling window estimation can accommodate variation in the predictability of characteristics, but it cannot be implemented in the IPCA framework.

In short, the fact that the rest of the 20 contract characteristics are not significant could be due to model misspecification or to the fact that the predictive power of these characteristics changes over time and the current version of the IPCA model cannot pick that up. This is supported by our empirical evidence that 12-month basis momentum is not significant in the full sample but is highly significant in the first sub-sample. We also have Fama-MacBeth cross-sectional regression results showing that only a small number of contract characteristics are significant over the full sample period. We use an alternative combination of a small set of contract characteristics in the cross-sectional regressions to take into account the small number (34) of commodity futures contracts available. This is in contrast to our results from Table 3 where 12 characteristics generate significant long-short returns.

7.2 Practical implications

The objective of asset pricing models is to identify pervasive risk factors, obtain estimates for betas, and construct expected returns. The estimated risk factors and expected returns are important inputs to asset allocation, performance evaluation, and corporate finance decisions.

Our evidence suggests that the performance of IPCA based models with latent risk factors is much better than the performance of traditional models with observable risk factors over the entire sample period from January 1981 to June 2022. Moreover, the performance of IPCA based latent factor models is much better during the second half of our sample period than during the first half of our sample period. The second half of our sample period corresponds to the post-financialization period of the commodity futures markets. Commodity futures returns have become much more correlated among themselves. However, some observable risk factors such as risk factors related to momentum, hedging pressure and speculative pressure, basis momentum, and idiosyncratic volatility still play significant roles in explaining cross-sectional variation in individual commodity futures contract returns. In practical applications, investors should take into account both latent factors and observable risk factors.

7.3 Risk-Based versus Behavior-Based Explanations

Prior studies tend to conclude that the observed profitability of momentum strategies in commodity futures markets cannot be explained by standard risk-factor models (Miffre and Rallis, 2007; Fuertes et al., 2010; Bianchi et al., 2015). Behavior-based explanations such as investor overreaction play at least a partial role (Shen et al., 2007; Bianchi et al., 2016).

Fuertes et al. (2015) conclude that profitability of trading strategies based on momentum, term structure (roll yields), and idiosyncratic volatility cannot be fully explained by a standard risk-factor model that includes both market and liquidity risk. They also conclude that behavioral explanations based on overreaction and subsequent mean reversion are unlikely to account for the performance of portfolio returns sorted on idiosyncratic volatility. Fernandez-Perez et al. (2016) find a significantly negative pricing relationship of idiosyncratic volatility in the commodity futures market. However, the pricing relationship vanishes when the fundamentals of backwardation and contango are suitably factored in the pricing model.

These earlier studies typically employ observable risk factors. They also do not allow betas to vary with contract characteristics. Our evidence from IPCA based latent factor models, on the other hand, suggests a strong role of risk-based explanation for both the momentum and idiosyncratic volatility effects. But we cannot rule out the role of behavior-based explanation as we do not carry out a formal test within the IPCA framework. It will be interesting to further explore these alternative explanations with the development of new statistical methods.

8. Conclusions

We introduce a dynamic factor-based asset pricing model in the commodity futures market. This is achieved by implementing the instrumented principal component analysis (IPCA). By estimating latent factors as functions of commodity futures characteristics, we show that a low-dimension factor model can successfully describe riskiness in commodity returns by explaining cross-sectional variations in average returns. Based on R^2's, our latent factor model outperforms a number of models using observable commodity risk factors, including models that also allow betas of observable factors to be dependent on commodity futures contract characteristics. We consider a total of seven widely-used risk factors in the commodity market related to market, momentum, basis, hedging pressure and speculative pressure, basis moment, idiosyncratic volatility, and skewness.

Only a small subset of commodity characteristics in the sample are responsible for IPCA's empirical success. Out of 20 commodity characteristics employed in the empirical analysis, an IPCA model with three characteristics is sufficient to describe returns. The three significant commodity market futures characteristics are momentum, expected shortfall, and idiosyncratic volatility. The tests show that these characteristics significantly contribute to the model by identifying latent factor loadings but show no statistical evidence of generating alphas. Therefore, the abnormal returns sorted on these variables are driven by compensation for higher exposures to latent factors rather than by mispricing not captured by latent factors. In addition, the three characteristics remain highly significant even when we allow seven observable factors to enter the IPCA model.

The key to the success of the model is *incorporating information from commodity characteristics into the estimation of factor loadings*. In the IPCA framework, risk loadings depend on observable commodity characteristics, which are treated as instrumental variables for estimating dynamic loadings on latent factors. The methodology improves the approach in the traditional dimension-reduction techniques, such as PCA, by allowing information beyond just returns into the estimation of factor loadings and alphas. Since factor loadings are dynamic and significant, we find strong evidence in favor of time-varying factor loadings and risk premiums in commodity futures returns.

References

Ahmed, S., and Tsvetanov, D., 2016. The predictive performance of commodity futures risk factors. Journal of Banking and Finance 71, 20-36.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31-56.

Asness, C.S., Moskowitz, T.J., and Pedersen, L.H., 2013. Value and momentum everywhere. Journal of Finance 68, 929-985.

Baba Yara, F., Boons, M., and Tamoni, A., 2019. Value return predictability across asset classes and commonalities in risk premia. Review of Financial Studies 25, 449-484.

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.

Barillas, F., and Shanken, J., 2017. Which alpha? The Review of Financial Studies 30, 1316-1338.

Bessembinder, H., and Chan, K., 1992. Time-varying risk premia and forecastable returns in futures markets. Journal of Financial Economics 32, 169-193.

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.

Berk, J.B., 2000. Sorting out sorts. Journal of Finance 55, 407-427.

Bianchi, R., Drew, M., and Fan, J. H., 2015. Combining momentum with reversal in commodity futures. Journal of Banking and Finance 59, 423-444.

Bianchi, R., Drew. M., and Fan, J. H., 2016. Commodities momentum: A behavioral perspective. Journal of Banking and Finance 72, 133-150.

Boons, M., and Prado, M.P., 2019. Basis momentum. Journal of Finance 74, 239-279.

Brogaard, J., Ringgenberg M., and Sovich D., 2019. The economic impact of index investing. Review of Financial Studies 32, 3461-3499.

Cheng, I., and Xiong W., 2014. Financialization of commodity markets. Annual Review of Financial Economics 6, 419-441.

Daniel, K., and Titman, S., 1997. Evidence on the characteristics of cross-sectional variation in stock returns. Journal of Finance 52, 1-33.

Daskalaki, C., Kostakis, A., and Skiadopoulos, G., 2014. Are there common factors in individual commodity futures returns? Journal of Banking and Finance 40, 346-363.

De Roon, F., Nijman T., and Veld, C., 2000. Hedging pressure effects in futures markets.

Journal of Finance 55, 1437-456.

Dewally, M., Ederington, L., and Fernando, C., 2013. Determinants of trader profits in commodity futures markets. Review of Financial Studies 26, 2648-2683.

Erb, C., and Harvey, C., 2006. The strategic and tactical value of commodity futures. Financial Analysts Journal 62, 69-97.

Fama, E., and French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

Fama, E., and French, K., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1-22.

Fama, E., and French, K., 2018. Choosing factors. Journal of Financial Economics 128, 234-252.

Fan, J., Fernandez-Perez, A., Fuertes, A., and Miffre, J. (2020). Speculative pressure. Journal of Futures Markets 40, 575-597.

Fan, J., Liao, Y., and Wang, W., 2016. Projected principal component analysis in factor models. Annals of Statistics 44, 219-254.

Fernandez-Perez, A., Frijns, B., Fuertes, A., and Miffre, J., 2019. A comprehensive appraisal of style-integration methods. Journal of Banking and Finance 105, 134-150.

Fernandez-Perez, A., Frijns, B., Fuertes, A., and Miffre, J., 2018. The skewness of commodity futures returns. Journal of Banking and Finance 86, 143-158.

Fernandez-Perez, A., Fuertes, A., and Miffre, J., 2016. Is idiosyncratic volatility priced in commodity futures markets? International Review of Financial Analysis 46, 219-226.

Ferson, W., and Harvey, C., 1997. Fundamental determinants of national equity market returns: A perspective on conditional asset pricing. Journal of Banking & Finance, 21, 1625-1665.

Fuertes, A., Miffre, J., and Rallis, G., 2010. Tactical allocation in commodity futures markets: Combining momentum and term structure signals. Journal of Banking and Finance 34, 2530-2548.

Fuertes, A., Miffre, J., and Fernandez-Perez, A., 2015. Commodity strategies based on momentum, term structure, and idiosyncratic volatility. Journal of Futures Markets 35, 274-297.

Fuertes, A., and Zhao, N., 2023. A Bayesian perspective on commodity style integration. Journal of Commodity Markets 30, 100328.

Gao, X., and Nardari, F., 2018. Do commodities add economic value in asset allocation? New evidence from time-varying moments. Journal of Financial and Quantitative Analysis 53, 365-393.

Gargano, A., and Timmermann, A., 2014. Forecasting commodity price indexes using macroeconomic and financial predictors. International Journal of Forecasting 30, 825-843.

Gorton, G., Hayashi, F., and Rouwenhorst, K. G., 2013. The fundamentals of commodity futures returns. Review of Finance 17, 35-105.

Gorton, G., and Rouwenhorst, K. G., 2006. Facts and fantasies about commodities futures. Financial Analysts Journal 62, 47-68.

Haugen, R., and Baker, N., 1996. Commonality in the determinants of expected stock returns. Journal of financial economics 41, 401-439.

Harvey, C. R., Liu, Y., and Zhu, H., 2016. ...and the cross-section of expected returns. The Review of Financial Studies 29, 5–68.

Henderson, B., Pearson, N., and Wang, L., 2015. New evidence on the financialization of commodity markets. Review of Financial Studies 28, 1285-1311.

Hong, H., and Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices? Journal of Financial Economics 105, 473-490.

Hou, K., Xue, C., and Zhang, L., 2015. Digesting anomalies: an investment approach. Review of Financial Studies 28, 650-705.

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

Kelly, B., Palhares, D., and Pruitt, S., 2023. Modeling corporate bond returns. Journal of Finance 78, 1967-2008.

Kelly, B., Pruitt, S., and Su, Y., 2019. Characteristics are covariances: a unified model of risk and return. Journal of Financial Economics 134, 501-524.

Kelly, B., Pruitt, S., and Su, Y., 2020. Instrumented principal component analysis. Technical Report.

Kim, S., Korajczyk, R., and Neuhierl, A., 2021. Arbitrage portfolios. Review of Financial Studies 34, 2813-2856.

Koijen, R., Moskowitz, T., Pedersen, L., and Vrugt, E., 2018. Carry. Journal of Financial Economics 127, 197-225.

Marshall, B., Cahan, R., and Cahan, J., 2008. Can commodity futures be profitably traded with quantitative market timing strategies? Journal of Banking and Finance 32, 1810–1819.

McLean, D.R., and Pontiff, J., 2016. Does academic research destroy return predictability? Journal of Finance 71, 5-32.

Miffre, J., 2016. Long-short commodity investing: a review of the literature. Journal of Commodity Markets 1, 3-13.

Miffre, J., and Rallis, G., 2007. Momentum strategies in commodity futures markets. Journal of Banking and Finance 31, 1863–1886.

Narayan, P., Ahmed, H., and Narayan, S., 2014. Do momentum-based trading strategies work in the commodity futures markets? Journal of Futures Markets 35, 868-891.

Sakkas, A., and Tessaromatis, N., 2020. Factor based commodity investing. Journal of Banking and Finance 115, 1-19.

Shen, Q., Szakmary, A., and Sharma, S., 2007. An examination of momentum strategies in commodity futures markets. Journal of Futures Markets 27, 227–256.

Subrahmanyam, A., 2010. The cross-section of expected stock returns: What have we learnt from the past twenty-five years of research? European Financial Management 16, 27-42.

Szymanowska, M., De Roon, F., Nijman, T., and Goorbergh, R., 2014. An anatomy of commodity futures risk premia. Journal of Finance 69, 453-482.

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

Wang, C., and Yu, M., 2004. Trading activity and price reversals in futures markets. Journal of Banking and Finance 28, 1337–1361.

Yang, F., 2013. Investment shocks and the commodity basis spread. Journal of Financial Economics 110, 164-184.

Table 1 Summary Statistics

The sample covers the period from January 1981 to June 2022. The table lists the 34 commodity futures included in this study. All commodity futures are classified into the following categories: energy, grains and oilseeds, livestock, industrial metals, and softs. The table reports summary statistics, including the starting month of each commodity futures contract, number of observations, mean, median, 5th percentile, 95th percentile, and standard deviation of annualized monthly returns. The last column reports the *t*-statistics to test the null hypothesis that the mean return is zero. Summary statistics are annualized. The last three rows provide summary statistics for the commodity market index return, the USD index return, and commodity produce price index (PPI) return. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Category	Category	Sample Starts	Sample Ends	Obs	Mean	5%	Median	95%	Std Dev.	Sharpe Ratio	*T*-stat.
Energy	Brent Crude Oil	198702	202206	425	0.045	-1.914	0.166	1.767	0.374	0.057	0.72
	Gas Oil	199101	202206	378	0.063	-1.888	0.129	1.696	0.310	0.125	1.14
	RBOB Gasoline	199802	202206	293	0.125	-1.689	0.264	1.646	0.308	0.327	2.00**
	Natural Gas	199007	202206	384	-0.108	-2.640	-0.123	2.336	0.440	-0.301	-1.39
	Heating Oil	198302	202206	473	0.075	-1.821	0.088	1.771	0.323	0.157	1.45
	WTI Crude Oil	199007	202206	384	0.075	-1.775	0.150	1.706	0.349	0.146	1.22
Grains and Oilseeds	Canola	199802	202206	293	0.035	-1.199	0.046	1.269	0.230	0.050	0.76
	Corn	198101	202206	498	-0.033	-1.521	-0.063	1.365	0.249	-0.228	-0.85
	Oats	199802	202206	293	0.057	-1.605	0.096	1.919	0.298	0.109	0.94
	Soybean Meal	199007	202206	384	0.091	-1.246	0.045	1.521	0.244	0.274	2.11**
	Soybean Oil	199007	202206	384	0.026	-1.297	-0.011	1.375	0.238	0.010	0.62
	Soybeans	198101	202206	498	0.043	-1.263	0.000	1.344	0.226	0.085	1.23
	Chicago Wheat	198101	202206	498	-0.043	-1.336	-0.018	1.397	0.258	-0.262	-1.08
	Kansas Wheat	199102	202206	377	-0.007	-1.519	-0.031	1.566	0.275	-0.114	-0.15
Livestock	Feeder Cattle	199101	202206	378	0.028	-0.817	0.072	0.853	0.148	0.024	1.05
	Lean Hogs	198101	202206	498	-0.027	-1.444	0.021	1.273	0.256	-0.199	-0.68
	Live Cattle	198101	202206	498	0.045	-0.820	0.054	0.822	0.144	0.144	2.00**
Metals	Aluminum	199102	202206	377	-0.013	-1.053	-0.051	1.141	0.191	-0.194	-0.38
	Copper	198101	202206	498	0.087	-1.180	0.074	1.532	0.251	0.252	2.24**
	Gold	198101	202206	498	0.024	-0.802	-0.003	0.925	0.162	0.002	0.97

	Palladium	199801	202206	294	0.120	-2.050	0.239	1.866	0.360	0.267	1.65
	Platinum	198401	202206	462	0.049	-1.179	0.055	1.155	0.223	0.110	1.35
	Silver	198101	202206	498	-0.003	-1.468	-0.050	1.668	0.285	-0.093	-0.06
	Lead	199101	202206	378	0.041	-1.332	0.087	1.620	0.264	0.065	0.87
	Nickel	199007	202206	384	0.060	-1.673	0.035	1.842	0.328	0.109	1.03
	Tin	199101	202206	378	0.079	-1.268	0.049	1.461	0.227	0.243	1.96
	Zinc	199102	202206	377	0.019	-1.201	-0.014	1.427	0.248	-0.022	0.42
Softs	Cocoa	198402	202206	461	-0.034	-1.527	-0.061	1.533	0.278	-0.209	-0.76
	Coffee	198102	202206	497	-0.010	-1.718	-0.103	2.072	0.348	-0.097	-0.18
	Cotton	198101	202206	498	0.025	-1.419	0.058	1.519	0.245	0.002	0.65
	Lumber	199802	202206	293	-0.013	-1.720	-0.077	1.981	0.338	-0.110	-0.19
	Orange Juice	199101	202206	378	-0.027	-1.808	-0.061	1.754	0.305	-0.167	-0.50
	Sugar	198101	202206	498	-0.036	-1.839	-0.017	1.768	0.345	-0.173	-0.67
	Rubber	199808	202206	287	-0.034	-1.722	-0.066	1.765	0.307	-0.190	-0.55
Commodity Market	Index Return	198101	202206	498	0.038	-1.207	0.069	1.078	0.205	0.071	1.21
USD	Index Return	198101	202206	498	0.031	-0.198	0.031	0.287	0.045	0.155	4.45***
Commodity PPI	Index Return	198101	202206	498	0.026	-0.154	0.024	0.215	0.034	0.070	5.03***

Table 2 Commodity Futures Characteristics and Commodity Futures Market Observable Risk Factors

The sample covers the period from January 1981 to June 2022. Panel A1 reports the portfolio returns sorted on the following 20 commodity futures contract characteristics: *MOM3*, *MOM12*, *CTR18*, *CTR36*, *R52WH*, *IVOL*, *SKEWNESS*, *MAX*, *ES*, *BASIS*, *HPHE*, *HPSP*, *BASM3*, *BASM12*, $BETA_{CMKT}$, $BETA_{INF}$, $BETA_{USD}$, *VOLM*, *OPNI*, and *ALIQ*. All characteristics are measured prior to month *t* when equally weighted portfolio returns are constructed. The summary statistics include the definition of each characteristic, mean, 5^{th} percentile, median, 95^{th} percentile, standard deviation, and number of observations. The summary statistics are calculated from pooled time-series and cross-sectional observations. The details of the construction of the sorting variables are provided in Appendix A. Panel A2 reports the pair-wise correlations among 20 commodity futures contract characteristics. Panel B1 reports the annualized summary statistics for commodity market observable risk factors *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12, CIVOL*, and *CSKEW*. Panel B2 reports the corresponding pairwise correlations among the seven observable risk factors. The number of observations in calculating pair-wise correlations can vary depending on the two variables under consideration *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A1: Summary Statistics for Commodity Futures Characteristics

	Definition	Mean	5%	Median	95%	Standard Deviation	Number of Observations
MOM3	3-month momentum return	0.007	-0.215	0.001	0.247	0.145	13944
MOM12	12-month momentum return	0.030	-0.426	-0.011	0.630	0.324	13800
CTR18	18-month contrarian return	0.043	-0.494	-0.023	0.805	0.411	13632
CTR36	36-month contrarian return	0.057	-0.596	-0.061	1.089	0.558	13416
R52WH	52-week high	0.853	0.560	0.889	1.000	0.147	13704
IVOL	Idiosyncratic volatility	0.014	0.006	0.013	0.024	0.007	13715
SKEWNESS	Skewness	-0.090	-0.882	-0.056	0.551	0.566	13715
MAX	Maximum daily in the past 12 months	0.058	0.024	0.051	0.109	0.051	13715
ES	Expected shortfall	-0.036	-0.064	-0.033	-0.018	0.016	13715
BASIS	Basis	0.005	-0.045	0.004	0.055	0.053	12617
HPHE	Hedging Pressure from Hedgers	-0.130	-0.521	-0.104	0.158	0.194	10088
HPSP	Hedging Pressure from Speculators	0.241	-0.246	0.255	0.683	0.294	10088
BASM3	3-month basis momentum	-0.000	-0.065	0.000	0.063	0.052	12589
BASM12	12-month basis momentum	-0.003	-0.070	-0.000	0.060	0.060	12505
$BETA_{CMKT}$	Commodity market beta	0.460	-0.011	0.288	1.518	0.482	13715
$BETA_{INF}$	Inflation beta	0.839	-2.236	0.724	4.780	2.254	13128
$BETA_{USD}$	USD index beta	-1.074	-3.968	-0.808	1.110	1.595	13128
VOLM	Average Daily Trading volume (million $)	2633.357	1.973	244.522	12281.384	8784.775	11452
OPNI	Average Daily Open interest (million $)	7190.182	3.239	1158.025	31209.133	18206.473	11556

ALIQ	Amihud liquidity measure	0.013	0.000	0.000	0.051	0.119	11427

Panel A2: Pairwise Correlations among Commodity Futures Contract Characteristics

	MOM3	*MOM12*	*CTR18*	*CTR36*	*R52WH*	*IVOL*	*SKEWNESSS*	*MAX*	*ES*	*BASIS*	*HPHE*	*HPSP*
MOM3												
MOM12	0.52***											
CTR18	0.40***	0.82***										
CTR36	0.26***	0.52***	0.66***									
R52WH	0.59***	0.68***	0.52***	0.32***								
IVOL	-0.01	-0.04***	-0.03***	-0.03***	-0.31***							
SKEWNESS	0.02**	-0.06***	-0.15***	-0.14***	0.03***	0.02**						
MAX	0.02**	-0.03***	-0.04***	-0.03***	-0.19***	0.73***	0.11***					
ES	0.04***	0.10***	0.05***	0.01	0.49***	-0.70***	0.27***	-0.60***				
BASIS	-0.13***	-0.23***	-0.27***	-0.26***	-0.17***	0.02*	0.06***	0.01	-0.02*			
HPHE	-0.11***	-0.29***	-0.30***	-0.20***	-0.20***	-0.08***	0.19***	-0.03***	0.08***	0.09***		
HPSP	0.14***	0.40***	0.41***	0.23***	0.27***	-0.02**	-0.19***	-0.03**	-0.03***	-0.12***	-0.75***	
BASM3	0.15***	0.04***	0.04***	0.01	0.06***	-0.01	0.00	-0.01	0.01	-0.30***	-0.02**	0.03**
BASM12	0.11***	0.16***	0.08***	0.00	0.14***	-0.05***	0.01	-0.04***	0.08***	-0.32***	-0.05***	0.08***
$BETA_{CMKT}$	0.01	0.04***	0.06***	0.06***	-0.17***	-0.15***	-0.17***	0.11***	-0.41***	-0.01	0.01	0.07***
$BETA_{INF}$	-0.02**	-0.07***	-0.09***	-0.13***	-0.17***	-0.08***	-0.09***	0.11***	-0.24***	0.05***	0.04***	-0.04***
$BETA_{USD}$	0.01	0.04***	0.06***	0.13***	0.19***	-0.08***	0.12***	-0.16***	0.35***	0.06***	0.09***	-0.14***
VOLM	0.03***	0.03***	0.01	-0.07***	-0.02**	-0.17***	-0.09***	0.10***	-0.14***	-0.01	0.01	0.07***
OPNI	0.04***	0.03***	0.00	-0.08***	-0.03***	-0.12***	-0.07***	0.12***	-0.16***	-0.01	0.05***	0.04***
ALIQ	-0.05***	-0.04***	-0.04***	-0.03***	-0.08***	-0.02**	-0.03***	0.05***	-0.07***	0.10***	-0.04***	0.00

	BASM3	*BASM12*	$BETA_{CMKT}$	$BETA_{INF}$	$BETA_{USD}$	*VOLM*	*OPNI*	*ALIQ*
MOM3								
MOM12								
CTR18								
CTR36								
R52WH								
IVOL								
SKEWNESS								
MAX								
ES								
BASIS								
HPHE								
HPSP								
BASM3								
BASM12	0.33***							
$BETA_{CMKT}$	0.00	-0.01						
$BETA_{INF}$	0.01	-0.01	0.44***					
$BETA_{USD}$	-0.01	0.00	-0.37***	-0.35***				
VOLM	0.01	0.01	0.35***	0.21***	-0.29***			
OPNI	0.01	0.01	0.35***	0.20***	-0.26***	0.85***		
ALIQ	0.00	0.01	0.07***	0.06***	-0.09***	-0.03***	-0.04***	

Panel B1: Summary Statistics for Commodity Futures Market Observable Risk Factors

	Definition	Mean	5%	Median	95%	Standard Deviation	Sharpe Ratio	Number of Observatio
CMKT	Commodity market-wide risk factor	0.038	-1.207	0.069	1.078	0.205	0.071	498
CMOM12	Momentum risk factor	0.107	-1.009	0.091	1.157	0.183	0.452	498
CBASIS	Basis risk factor	0.062	-0.980	0.083	1.010	0.177	0.215	498
CHP	Hedging pressure risk factor	0.078	-1.012	0.090	1.095	0.187	0.288	426
CBAS12	Basis momentum risk factor	0.054	-0.902	0.035	1.050	0.180	0.167	498
CIVOL	Idiosyncratic volatility risk factor	0.093	-0.862	0.083	1.056	0.178	0.389	498
CSKEW	Skewness risk factor	0.068	-0.850	0.066	0.986	0.161	0.275	498

Panel B2: Pairwise Correlations between Commodity Futures Market Observable Risk Factors

	CMKT	*CMOM12*	*CBASIS*	*CHP*	*CBASM12*	*CIVOL*
CMOM12	0.11**					
CBASIS	0.21***	0.29***				
CHP	0.00	0.22***	0.06			
CBASM12	0.03	0.25***	0.47***	0.07		
CIVOL	0.28***	0.15***	0.22***	-0.14***	0.10**	
CSKEW	0.09**	0.06	0.17***	0.23***	0.03	-0.03

Table 3 Hedge Portfolio Returns Sorted on Commodity Futures Characteristics

The sample covers the period from January 1981 to June 2022. The table reports the portfolio returns sorted on the following 20 commodity futures characteristics: *MOM3*, *MOM12*, *CTR18*, *CTR36*, *R52WH*, *IVOL*, *SKEWNESS*, *MAX*, *ES*, *BASIS*, *HPHE*, *HPSP*, *BASM3*, *BASM12*, $BETA_{CMKT}$, $BETA_{INF}$, $BETA_{USD}$, *VOLM*, *OPNI*, and *ALIQ*. All characteristics are measured prior to month *t* when equally weighted portfolio returns are constructed. The sorting is based on the 30% and 70% value of each characteristic in each month *t-1*. The summary statistics include the mean returns on the long position, the mean returns on the short position, the returns on the hedge portfolios (long – short), *t*-statistics, and number of monthly observations. The reported returns on long, short, and long-short positions are annualized. The *t*-statistic tests for the null hypothesis that the mean return from the long-short portfolio is zero. The details of the construction of the sorting variables are provided in Appendix A. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Sorting Characteristic	Long	Short	Long - Short	*t*-statistic	Number of Months
MOM3	0.083	-0.016	0.099	3.25***	498
MOM12	0.078	-0.029	0.107	3.75***	498
CTR18	0.039	-0.011	0.050	1.80*	498
CTR36	0.012	0.009	0.003	0.11	498
R52WH	0.077	-0.037	0.114	4.28***	498
IVOL	-0.030	0.063	-0.093	-3.37***	498
SKEWNESS	-0.027	0.041	-0.068	-2.73***	498
MAX	-0.033	0.043	-0.076	-2.68***	498
ES	0.044	0.008	0.037	1.36	498
BASIS	-0.004	0.058	-0.062	-2.26**	498
HPHE	0.004	0.060	-0.056	-2.05**	426
HPSP	0.054	-0.011	0.065	2.41**	426
BASM3	0.013	0.040	-0.027	-1.00	498
BASM12	0.050	-0.004	0.054	1.94*	498
$BETA_{CMKT}$	0.021	0.024	-0.003	-0.11	498
$BETA_{INF}$	0.025	0.032	-0.007	-0.24	498
$BETA_{USD}$	0.007	0.006	0.001	0.03	498
VOLM	-0.036	0.027	-0.063	-2.56**	498
OPNI	-0.034	0.044	-0.079	-3.54***	498
ALIQ	0.040	-0.018	0.058	2.51**	498

Table 4 Estimate of the Basic IPCA Model With and Without the Intercept

The table estimates the IPCA model using monthly returns from 34 commodity futures contracts. The number of latent factors (*K*) is equal to 1, 2, and 3. The estimation in Panel A employs four commodity futures characteristics: Z=(*MOM12*, *CTR36*, *ES*, *IVOL*). The sample covers the period from January 1981 to June 2022, for a total of 498 months. The estimation in Panel B employs seven commodity futures characteristics: Z=(*MOM12*, *CTR36*, *ES*, *IVOL*, *BASIS*, *BASM12*, *ALIQ*). The sample covers the period from February 1984 to June 2022, for a total of 461 months. All characteristics are measured in month *t-1* prior to month *t* when commodity futures returns are measured. The details of the construction of the commodity futures contract characteristics are provided in Appendix A. The table reports the R^2's in percentages for the restricted model when alpha is equal to 0 ($\Gamma_\alpha = 0$) and the unrestricted model when alpha is not equal to 0 ($\Gamma_\alpha \neq 0$). The table also reports the bootstrapped *p*-values for testing the null hypothesis that $\Gamma_\alpha = 0$ under the unrestricted model. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Number of commodity futures contract characteristics L = 4

		1	2	3
R^2				
	$\Gamma_\alpha = 0$	0.254	0.339	0.393
	$\Gamma_\alpha \neq 0$	0.257	0.340	0.394
W_α *p*-value for testing $\Gamma_\alpha = 0$		0.005***	0.027**	0.806

Panel B: Number of commodity futures contract characteristics L = 7

		1	2	3
R^2				
	$\Gamma_\alpha = 0$	0.262	0.346	0.410
	$\Gamma_\alpha \neq 0$	0.266	0.348	0.411
W_α *p*-value for testing $\Gamma_\alpha \neq 0$		0.004***	0.004***	0.647

Table 5 The Significance of Commodity Futures Characteristics in the IPCA Model

The table estimates the IPCA model using monthly returns from 34 commodity futures contracts. The number of latent factors (*K*) is equal to 3. The baseline IPCA model in Panel A employs four commodity futures characteristics (L = 4): Z = (*MOM12*, *CTR36*, *ES*, *IVOL*). Then Panel A estimates the IPCA model by adding each of the remaining 16 characteristics, one at a time. The sample covers the period from January 1981 to June 2022, for a total of 498 months. The baseline IPCA model in Panel B employs seven commodity futures characteristics (L = 7): Z = (*MOM12*, *CTR36*, *ES*, *IVOL*, *BASIS*, *BASM12*, *ALIQ*). Then Panel B estimates the IPCA model by adding each of the remaining 13 characteristics, one at a time. The sample covers the period from February 1984 to June 2022, for a total of 461 months. All characteristics are measured in month *t-1* prior to month *t* when commodity futures returns are measured. The details of the construction of commodity futures contract characteristics are provided in Appendix A. The IPCA models are restricted when alpha is equal to 0 ($\Gamma_\alpha = 0$). The table reports the bootstrapped *p*-values for testing the null hypothesis when the coefficient associated with the particular characteristic in question is zero. In all model specifications, betas of all latent factors are instrumented with commodity futures characteristics. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: IPCA Model with *K*=3 and L=4

First 4 Characteristics				5th Characteristics	
MOM12	*CTR36*	*ES*	*IVOL*		
0.000***	0.742	0.003***	0.000***		
0.718	0.767	0.010***	0.000***	*MOM3*	0.002***
0.000***	0.883	0.007***	0.000***	*CTR18*	0.599
0.097*	0.547	0.006***	0.000***	*R52WH*	0.695
0.000***	0.716	0.017**	0.000***	*SKEWNESS*	0.743
0.000***	0.714	0.060*	0.000***	*MAX*	0.195
0.000***	0.840	0.000***	0.000***	*HPHE*	0.672
0.000***	0.854	0.002***	0.000***	*HPSP*	0.847
0.000***	0.939	0.015**	0.000***	*BASIS*	0.649
0.000***	0.897	0.022**	0.002***	*BASM3*	0.980
0.000***	0.913	0.007***	0.000***	*BASM12*	0.942
0.000***	0.785	0.126	0.003***	$BETA_{CMKT}$	0.128
0.000***	0.605	0.026**	0.000***	$BETA_{INF}$	0.778
0.000***	0.691	0.021**	0.001***	$BETA_{USD}$	0.691
0.000***	0.853	0.021**	0.000***	*VOLM*	0.251
0.000***	0.851	0.022**	0.001***	*OPNI*	0.257
0.000***	0.860	0.021**	0.000***	*ALIQ*	0.557

Panel B: IPCA Model with *K*=3 and L=7

First 7 Characteristics							8th Characteristic	
MOM12	*CTR36*	*ES*	*IVOL*	*BASIS*	*BASM12*	*ALIQ*		
0.000***	0.796	0.031**	0.003***	0.555	0.870	0.437		
0.177	0.769	0.011**	0.001***	0.677	0.962	0.393	*MOM3*	0.157
0.000***	0.825	0.007***	0.003***	0.485	0.831	0.455	*CTR18*	0.815
0.000***	0.690	0.007***	0.004***	0.403	0.848	0.388	*R52WH*	0.862
0.000***	0.745	0.014**	0.013**	0.567	0.870	0.609	*SKEWNESS*	0.827
0.000***	0.624	0.126	0.001***	0.501	0.831	0.492	*MAX*	0.416
0.000***	0.748	0.029**	0.003***	0.625	0.901	0.524	*HPHE*	0.582
0.000***	0.726	0.025**	0.003***	0.625	0.915	0.701	*HPSP*	0.872
0.000***	0.770	0.008***	0.005***	0.557	0.927	0.477	*BASM3*	0.920
0.000***	0.801	0.237	0.002***	0.508	0.860	0.976	$BETA_{CMKT}$	0.330
0.000***	0.762	0.035**	0.000***	0.527	0.868	0.439	$BETA_{INF}$	0.736
0.004***	0.740	0.039**	0.000***	0.508	0.850	0.375	$BETA_{USD}$	0.589
0.000***	0.761	0.019**	0.009***	0.495	0.865	0.486	*VOLM*	0.502
0.000***	0.645	0.010***	0.004***	0.499	0.851	0.271	*OPNI*	0.262

Table 6 The IPCA Models with Observable Factors

The table estimates the IPCA model using 34 commodity futures contracts, controlling for seven observable risk factors in the commodities markets. These include risk factors related to market (*CMKT*), momentum (*CMOM12*), basis (*CBASIS*), hedging-speculative pressure (*CHP*), basis momentum (*CBASM12*), idiosyncratic volatility (*CIVOL*), and skewness (*CSKEW*) in the commodities markets. The table summarizes R^2's from alternative model specifications. The number of latent factors (*K*) is equal to 1, 2, and 3. The sample covers the period from January 1987 to June 2022, for a total of 426 months. Panel A employs four commodity futures characteristics (L=4): Z=(*MOM12*, *CTR36*, *ES*, *IVOL*). Panel B employs seven commodity futures contract characteristics (L=7): Z=(*MOM12*, *CTR36*, *ES*, *IVOL*, *BASIS*, *BASM12*, *ALIQ*). All characteristics are measured in month *t-1* prior to month *t* when commodity futures returns are measured. The details of the construction of commodity futures contract characteristics are provided in Appendix A. The IPCA models are restricted when alpha is equal to 0 (($\Gamma_\alpha = 0$). The table also reports the bootstrapped *p*-values for testing the null hypothesis when the coefficients associated with each of the observable risk factors are zero. In all model specifications, both betas to latent and observable factors are instrumented with commodity futures characteristics. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: Number of commodity futures contract characteristics L = 4

	Number of Latent Factors *K*		
	1	2	3
IPCA + 7 Observable Factors			
R^2			
Latent factors only	0.254	0.339	0.393
Latent factors + *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*	0.390	0.418	0.438
Bootstrapped p-value			
CMKT	0.010***	0.000***	0.261
CMOM12	0.000***	0.000***	0.000***
CBASIS	0.000***	0.392	0.211
CHP	0.002***	0.000***	0.000***
CBASM12	0.000***	0.952	0.709
CIVOL	0.010***	0.000***	0.000***
CSKEW	0.051*	0.046**	0.243

Panel B: Number of commodity futures contract characteristics L = 7

	Number of Latent Factors *K*		
	1	2	3
IPCA + 7 Observable Factors			
R^2			
Latent factors only	0.262	0.346	0.410
Latent factors + *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*	0.437	0.469	0.491
Bootstrapped p-value			
CMKT	0.001***	0.000***	0.000***
CMOM12	0.000***	0.000***	0.659
CBASIS	0.000***	0.000***	0.000***
CHP	0.001***	0.004***	0.000***
CBASM12	0.000***	0.000***	0.000***
CIVOL	0.069*	0.000***	0.000***
CSKEW	0.200	0.071*	0.524

Table 7 Observable Factors and the Significance of Commodity Futures Characteristics

The sample covers the period from January 1987 to June 2022, for a total of 426 months. The table tests for the significance of commodity futures characteristics when observable factors are also included in the IPCA model. The table estimates the IPCA model using monthly returns from 34 commodity futures contracts. The number of latent factors is equal to three, i.e., *K*=3. In all model specifications, betas of all latent factors are instrumented with three commodity futures characteristics (L=3): Z=(*MOM12, ES, IVOL*). The table estimates the IPCA model, adding varying combinations of observable risk factors *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*. All characteristics are measured in month *t-1* prior to month *t* when commodity futures returns are measured. The details of the construction of commodity futures contract characteristics are provided in Appendix A. The IPCA models are restricted when alpha is equal to 0 (($\Gamma_\alpha = 0$). The table reports the bootstrapped *p*-values for testing the null hypothesis when the coefficient associated with a particular characteristic is zero. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

IPCA Model with *K*=3, L=3 and Observable Factors				
Model	*MOM12*	*ES*	*IVOL*	Observable Factors Included in the IPCA Model
1	0.000***	0.000***	0.000***	*CMKT*
2	0.271	0.000***	0.000***	*CMKT*, *CMOM12*
3	0.282	0.000***	0.000***	*CMKT*, *CMOM12*, *CBASIS*
4	0.476	0.000***	0.000***	*CMKT*, *CMOM12*, *CBASIS*, *CHP*
5	0.461	0.000***	0.000***	*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*
6	0.000***	0.000***	0.018**	*CMKT*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*
7	0.000***	0.000***	0.006***	*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*
8	0.000***	0.000***	0.008***	*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*

Table 8 Sub-Sample IPCA Estimation Results

The table estimates the IPCA model using monthly returns from 34 commodity futures contracts. The full sample is divided into two sub-samples. The first sub-sample covers the period from January 1981 to December 2002, for a total of 264 months. The second sub-sample covers the period from January 2003 to June 2022, for a total of 234 months. The baseline IPCA model in Panel A employs three commodity futures characteristics (L = 3): *Z* = (*MOM12*, *ES*, *IVOL*). Panel A compares R^2's from IPCA models with *K* = 1, 2, and 3, respectively. Panel B tests for the significance of individual contract characteristics for the first sub-sample. Panel C tests for the significance of individual contract characteristics for the second sub-sample. The number of latent factors in Panels B and C is equal to three (*K* = 3). All characteristics are measured in month *t-1* prior to month *t* when commodity futures returns are measured. The details of the construction of commodity futures contract characteristics are provided in Appendix A. The IPCA models are restricted when alpha is equal to 0 ($\Gamma_\alpha = 0$). The table reports the bootstrapped *p*-values for testing the null hypothesis when the coefficient associated with the particular characteristic in question is zero. In all model specifications, betas of all latent factors are instrumented with commodity futures characteristics. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: The Performance of the IPCA Model during Two Sub-Samples

	Number of Observations	R^2		
		K=1	*K*=2	*K*=3
First half sub-sample from January 1981 to December 2002				
IPCA with *Z* only	5,759	0.180	0.267	0.335
Second half sub-sample from January 2003 to June 2022				
IPCA with *Z* only	7,956	0.300	0.381	0.421

Panel B: Significance of Individual Characteristics, First Half Sub-Sample from January 1981 to December 2002

First 3 Characteristics			4th Characteristics	
MOM12	*ES*	*IVOL*		
0.000***	0.001***	0.000***		
0.981	0.002***	0.005***	*MOM3*	0.000***
0.000***	0.002***	0.001***	*CTR18*	0.556
0.004***	0.000***	0.000***	*CTR36*	0.504
0.000***	0.007***	0.002***	*R52WH*	0.885
0.000***	0.001***	0.000***	*SKEWNESS*	0.867
0.000***	0.005***	0.000***	*MAX*	0.329
0.000***	0.000***	0.000***	*HPHE*	0.611
0.000***	0.000***	0.000***	*HPSP*	0.815
0.000***	0.013**	0.001***	*BASIS*	0.727
0.000***	0.010**	0.033**	*BASM3*	0.075*
0.000***	0.015**	0.038**	*BASM12*	0.000***
0.000***	0.010**	0.001***	$BETA_{CMKT}$	0.029**
0.000***	0.002***	0.000***	$BETA_{INF}$	0.930
0.000***	0.004***	0.002***	$BETA_{USD}$	0.760
0.000***	0.046**	0.001***	*VOLM*	0.313
0.000***	0.026**	0.001***	*OPNI*	0.301
0.000***	0.022**	0.000***	*ALIQ*	0.638

Panel C: Significance of Individual Characteristics, Second Half Sub-Sample from January 2003 to June 2022

First 3 Characteristics			4^{th} Characteristics	
MOM12	*ES*	*IVOL*		
0.000***	0.002***	0.000***		
0.661	0.003***	0.000***	*MOM3*	0.000***
0.000***	0.014**	0.000***	*CTR18*	0.170
0.000***	0.021**	0.000***	*CTR36*	0.195
0.927	0.001***	0.000***	*R52WH*	0.000***
0.000***	0.008***	0.000***	*SKEWNESS*	0.787
0.000***	0.078*	0.000***	*MAX*	0.168
0.000***	0.003***	0.000***	*HPHE*	0.946
0.000***	0.004***	0.000***	*HPSP*	0.666
0.000***	0.010**	0.000***	*BASIS*	0.287
0.000***	0.002***	0.000***	*BASM3*	0.753
0.000***	0.004***	0.000***	*BASM12*	0.687
0.000***	0.102	0.001***	$BETA_{CMKT}$	0.136
0.000***	0.016**	0.000***	$BETA_{INF}$	0.705
0.000***	0.105	0.000***	$BETA_{USD}$	0.744
0.000***	0.001***	0.000***	*VOLM*	0.381
0.000***	0.004***	0.000***	*OPNI*	0.471
0.000***	0.004***	0.000***	*ALIQ*	0.475

Table 9 Compare Alternative Asset Pricing Models

The table compares R^2's from alternative asset pricing models using 34 commodity market futures contracts. The sample period is from January 1987 to June 2022. Panel A summarizes R^2's from (a) latent factor models from IPCA with betas instrumented by four commodity futures contract characteristics (L=4), with and without observable risk factors; and (b) latent factor models from IPCA with betas instrumented by seven commodity futures contract characteristics (L=7), with and without observable risk factors. The four commodity futures contract characteristics are *MOM12*, *CTR36*, *ES*, and *IVOL*. The seven commodity futures contract characteristics are *MOM12*, *CTR36*, *ES*, *IVOL*, *BASIS*, *BASM12*, and *ALIQ*. The number of latent factors (*K*) is equal to 1, 2, and 3. Panel B reports R^2 from (a) observable risk factor models, with and without beta being instrumented by four commodity futures contract characteristics (L=4); and (b) observable risk factor models, with and without beta being instrumented by seven commodity futures contract characteristics (L=7). The three-factor model includes the market, momentum, and basis factors (*CMKT*, *CMOM12*, and *CBASIS*). The four-factor model includes the market, momentum, basis, and hedging-speculative pressure factors (*CMKT*, *CMOM12*, *CBASIS*, and *CHP*). The five-factor model includes the market, momentum, basis, hedging-speculative pressure, and basis momentum factors (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, and *CBASM12*). The six-factor model includes the market, momentum, basis, hedging-speculative pressure, basis momentum, and idiosyncratic volatility (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, and *CIVOL*). The seven-factor model includes the market, momentum, basis, hedging and speculative pressure, basis momentum, idiosyncratic volatility, and skewness factors (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*. All characteristics are measured in month *t-1* prior to month *t* when commodity futures contract returns are measured. The details of the construction of commodity futures contract characteristics are provided in Appendix A.

Panel A: R^2's from IPCA Models

	Number of Observations	R^2		
		K=1	*K*=2	*K*=3
L = 4				
IPCA with *Z* only	12,603	0.254	0.339	0.393
IPCA with *Z* and 3-factor model (*CMKT*, *CMOM12*, *CBASIS*) with Z	12,603	0.363	0.397	0.425
IPCA with *Z* and 4-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*) with Z	12,603	0.365	0.399	0.427
IPCA with *Z* and 5-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*) with Z	12,603	0.367	0.399	0.427
IPCA with *Z* and 6-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*) with Z	12,603	0.388	0.417	0.438
IPCA with *Z* and 7-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*) with *Z*	12,603	0.390	0.418	0.438
L = 7				
IPCA with *Z* only	10,150	0.262	0.346	0.410
IPCA with *Z* and 3-factor model (*CMKT*, *CMOM12*, *CBASIS*) with Z	10,150	0.395	0.433	0.466
IPCA with *Z* and 4-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*) with Z	10,150	0.399	0.435	0.469
IPCA with *Z* and 5-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*) with Z	10,150	0.419	0.453	0.485
IPCA with *Z* and 6-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*) with Z	10,150	0.435	0.468	0.491
IPCA with *Z* and 7-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*) with *Z*	10,150	0.437	0.469	0.491

Panel B: R^2's from Models Based on Observable Risk Factors

	Number of Observations	R^2
L= 4		
Observable factors without *Z*		
3-factor model (*CMKT*, *CMOM12*, *CBASIS*)	12,603	0.143
4-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*)	12,603	0.148
5-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*)	12,603	0.149
6-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*)	12,603	0.169
7-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*)	12,603	0.169
Observable factors with Z		
3-factor model (*CMKT*, *CMOM12*, *CBASIS*)	12,603	0.260
4-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*)	12,603	0.268
5-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*)	12,603	0.270
6-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*)	12,603	0.322
7-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*)	12,603	0.325
L= 7		
Observable factors without *Z*		
3-factor model (*CMKT*, *CMOM12*, *CBASIS*)	10,150	0.150
4-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*)	10,150	0.154
5-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*)	10,150	0.156
6-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*)	10,150	0.174
7-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*)	10,150	0.175
Observable factors with *Z*		
3-factor model (*CMKT*, *CMOM12*, *CBASIS*)	10,150	0.299
4-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*)	10,150	0.307
5-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*)	10,150	0.330
6-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*)	10,150	0.376
7-factor model (*CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*)	10,150	0.378

Table 10 Spanning Tests and Sharpe Ratios

The sample covers the period from January 1987 to June 2022. The table constructs latent factors *FAC1*, *FAC2*, and *FAC3* from the IPCA model using 34 commodity futures contracts. The number of commodity futures contract characteristics are four (L=4) and seven (L=7), respectively. The number of latent factors (*K*) is equal to 3. Panels A1 and A2 implement spanning tests when the dependent variables are IPCA latent factors *FAC1*, *FAC2*, and *FAC3*, respectively. The independent variables are observable risk factors *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*. Panels B1 and B2 implement spanning tests when the dependent variables are observable risk factors *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*, respectively. The independent variables are IPCA latent factors *FAC1*, *FAC2*, and *FAC3*. Panel C reports the annualized Sharpe ratios for individual risk factors and annualized Sharpe ratios for tangent portfolios, including tangent portfolios from the IPCA latent factors, commodity market observable risk factors, and Fama and French (2015) equity market risk factors *EXMRET*, *SMB*, *HML*, *RMW*, and *CMA* plus the momentum factor *MOM*. The risk-free rate is 2.4% per annum. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A1: The Dependent Variables are Latent Factors (L=4)

	CONST	*CMKT*	*CMOM12*	*CBASIS*	*CHP*	*CBASM*	*CIVOL*	*CSKEW*	*R2*	*Obs.*
FAC1	0.257	1.355***	-0.426***	0.150	-0.179	-0.008	1.497***	0.534***		
	(0.55)	(11.52)	(-3.32)	(1.05)	(-1.61)	(-0.05)	(12.07)	(4.23)	0.652	426
FAC2	0.219	0.054	1.216***	-0.305***	0.217***	0.133	0.274***	-0.173**		
	(0.85)	(1.06)	(18.75)	(-4.06)	(3.80)	(1.46)	(4.32)	(-2.58)	0.640	426
FAC3	0.198*	0.625***	-0.053*	-0.044	0.105***	0.058**	-0.475***	-0.023		
	(1.70)	(21.22)	(-1.80)	(-1.23)	(3.83)	(2.07)	(-15.74)	(-0.73)	0.727	426

Panel A2: The Dependent Variables are Latent Factors (L=7)

	CONST	*CMKT*	*CMOM12*	*CBASIS*	*CHP*	*CBASM*	*CIVOL*	*CSKEW*	*R2*	*Obs.*
FAC1	0.146	1.231***	-0.312**	0.072	0.019	0.411**	1.352***	0.539***		
	(0.29)	(13.87)	(-2.50)	(0.40)	(0.13)	(2.13)	(10.11)	(4.49)	0.597	426
FAC2	-0.012	-0.075	1.320***	-0.429***	0.251***	0.242**	-0.001	-0.250***		
	(-0.04)	(-1.25)	(16.04)	(-4.36)	(3.93)	(2.09)	(-0.02)	(-3.12)	0.569	426
FAC3	-0.139	-0.641***	0.057*	0.088**	-0.102***	-0.046	0.521***	0.048		
	(-1.04)	(-22.26)	(1.71)	(2.02)	(-3.44)	(-1.29)	(16.35)	(1.48)	0.686	426

Panel B1: The Dependent Variables are Pre-specified Factors (L=4)

	CONST	*FAC1*	*FAC2*	*FAC3*	*R2*	*Obs.*
CMKT	-0.302**	0.231***	0.109***	0.856***		
	(-2.57)	(14.79)	(5.30)	(21.06)	0.843	426
CMOM12	0.205	0.002	0.470***	0.025		
	(1.20)	(0.14)	(19.46)	(0.60)	0.578	426
CBASIS	0.341	0.069***	0.067*	0.047		
	(1.57)	(3.44)	(1.93)	(0.80)	0.072	426
CHP	0.480*	-0.043*	0.161***	0.230***		
	(1.82)	(-1.81)	(4.71)	(3.10)	0.113	426
CBASM12	0.297	0.008	0.140***	0.008		
	(1.48)	(0.41)	(3.58)	(0.16)	0.068	426
CIVOL	0.175	0.187***	0.114***	-0.311***		
	(1.10)	(11.29)	(4.55)	(-6.34)	0.502	426
CSKEW	0.748***	0.039**	0.003	0.047		
	(3.39)	(2.03)	(0.10)	(0.73)	0.024	426

Panel B2: The Dependent Variables are Pre-specified Factors (L=7)

	CONST	*FAC1*	*FAC2*	*FAC3*	*R2*	*Obs.*
CMKT	-0.138	0.246***	0.036	-0.801***		
	(-0.95)	(15.75)	(1.60)	(-20.67)	0.777	426
CMOM12	0.399**	0.043**	0.386***	-0.024		
	(2.21)	(2.24)	(16.69)	(-0.64)	0.506	426
CBASIS	0.373*	0.085***	0.025	0.003		
	(1.71)	(3.98)	(0.79)	(0.05)	0.077	426
CHP	0.520**	-0.004	0.154***	-0.192***		
	(2.01)	(-0.17)	(4.62)	(-2.89)	0.101	426
CBASM12	0.286	0.050***	0.131***	0.025		
	(1.41)	(2.80)	(3.81)	(0.55)	0.096	426
CIVOL	0.249	0.179***	0.019	0.274***		
	(1.49)	(11.41)	(0.72)	(6.81)	0.416	426
CSKEW	0.742***	0.049**	-0.008	-0.018		
	(3.37)	(2.55)	(-0.30)	(-0.31)	0.031	426

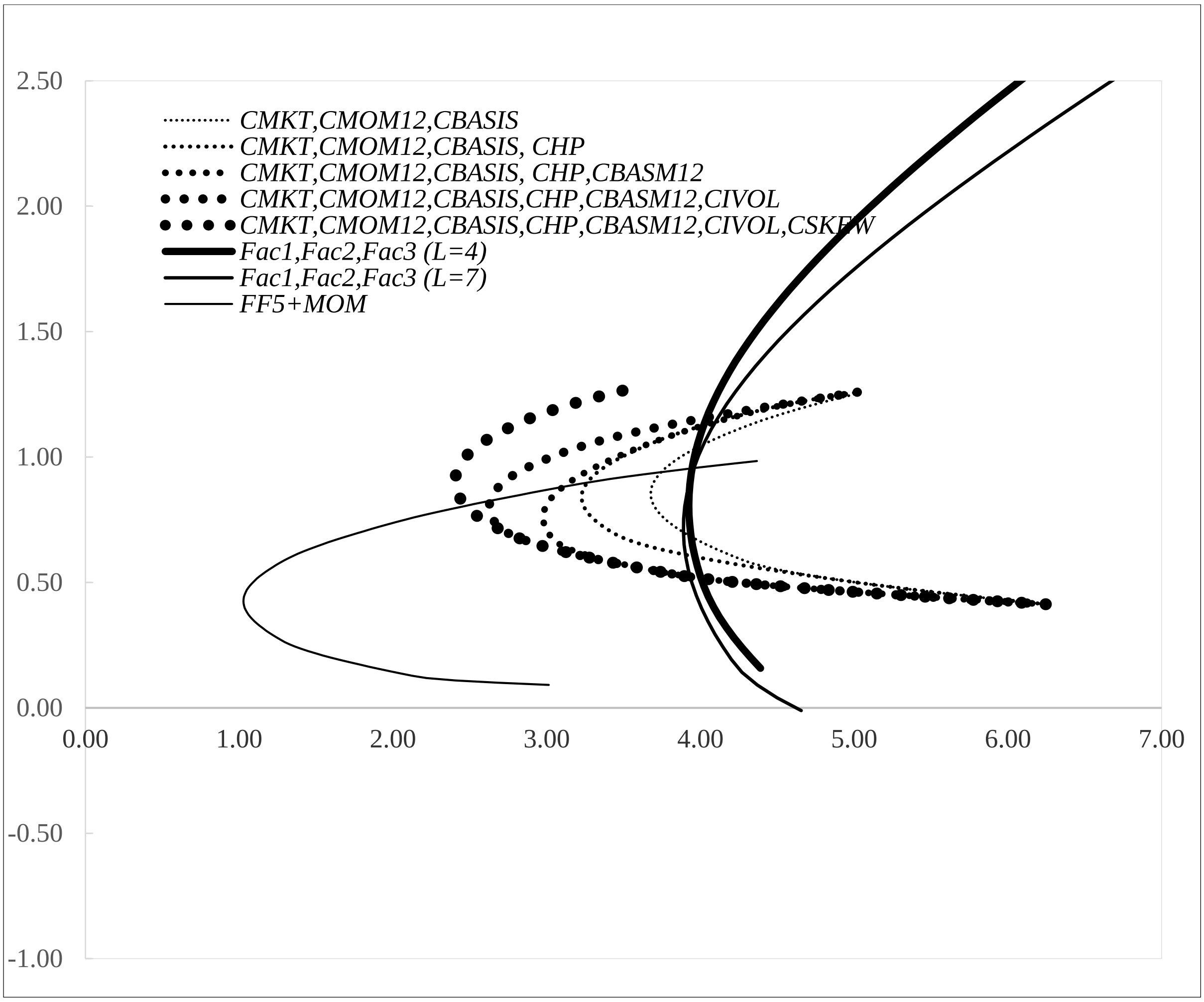
Panel C: Sharpe Ratios of Individual Portfolios and Tangent Portfolios

	L=4	L=7		
FAC1	0.310	0.358		
FAC2	0.457	0.250		
FAC3	-0.042	-0.157		
FAC1, *FAC2*	0.565	0.464		
FAC1, *FAC2*, *FAC3*	0.565	0.464		
CMKT			0.081	
CMOM12			0.419	
CBASIS			0.249	
CHP			0.288	
CBASM12			0.226	
CIVOL			0.250	
CSKEW			0.435	
CMKT, *CMOM12*			0.421	
CMKT, *CMOM12*, *CBASIS*			0.439	
CMKT, *CMOM12*, *CBASIS*, *CHP*			0.482	
CMKT, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*			0.485	
CMKT, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*			0.541	
CMKT, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*			0.649	
FAC1, *FAC2*, *FAC3*, *CMKT*, *CMOM12*, *CBASIS*	0.583	0.542		
FAC1, *FAC2*, *FAC3*, *CMKT*, *CMOM12*, *CBASIS*, *CHP*	0.618	0.583		
FAC1, *FAC2*, *FAC3*, *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*	0.620	0.584		
FAC1, *FAC2*, *FAC3*, *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*	0.622	0.597		
FAC1, *FAC2*, *FAC3*, *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, *CSKEW*	0.737	0.703		
EXMRET				0.377
SMB				-0.129
HML				-0.047
RMW				0.234
CMA				0.102
MOM				0.219

EXMRET, *SMB*, *HML*	0.379
EXMRET, *SMB*, *HML*, *MOM*	0.503
EXMRET, *SMB*, *HML*, *RMW*, *CMA*	0.574
EXMRET, *SMB*, *HML*, *RMW*, *CMA*, *MOM*	0.668

Figure 1: Efficient Frontiers

Panel A of the figure plots the efficient frontier constructed from the following assets: (i) *CMKT*, *CMOM12*, and *CBASIS*; (ii) *CMKT*, *CMOM12*, *CBASIS*, and *CHP* ; (iii) *CMKT*, *CMOM12*, *CBASIS*, *CHP*, and *CBASM12*; (iv) *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, and *CIVOL*; (v) *CMKT*, *CMOM12*, *CBASIS*, *CHP*, *CBASM12*, *CIVOL*, and *CSKEW*; (vi) *FAC1*, *FAC2*, and *FAC3* (L=4); (vii) *FAC1*, *FAC2*, and *FAC3* (L=7); and (viii) *FF5* (*EXMRET*, *SMB*, *HML*, *RMW*, and *CMA*) plus *MOM*. The mean and standard deviations of the returns from these assets are computed for the period January 1987 to June 2022.



Appendix A: Construction of 20 Commodity Futures Contract Characteristics

Appendix A provides details of the construction of 20 characteristics of commodity futures contracts: *MOM3*, *MOM12*, *CTR18*, *CTR36*, *R52WH*, *IVOL*, *SKEWNESS*, *MAX*, *ES*, *BASIS*, *HPHE*, *HPSP*, *BASM3*, *BASM12*, $BETA_{CMKT}$, $BETA_{INF}$, $BETA_{USD}$, *VOLM*, *OPNI*, and *ALIQ*. All characteristics are measured in month *t-1* prior to month *t* when equal weighted portfolio returns are constructed.

Variable Names	Details of Construction
3-month momentum measure (*MOM3*)	*MOM3* = cumulative return over prior months *t-3* to *t-1*
12-month momentum measure (*MOM12*)	*MOM12* = cumulative return over prior months *t-12* to *t-1*
18-month contrarian measure (*CTR18*)	*CTR18* = cumulative return over month *t-18* to month *t-1*
36-month contrarian measure (*CTR36*)	*CTR36* = cumulative return over month *t-36* to month *t-1*
52-week high (*R52WH*)	The ratio of highest price in the past 12 months to the monthly price in the past 12 months. The price is computed by cumulative monthly returns of commodity futures contracts from *t-12* to *t-1*.
Idiosyncratic volatility (*IVOL*)	Residual standard deviation from the following regression: $r_{i,t} = \beta_0 + \beta_1 r_{m,t} + \varepsilon_{i,t}$, where $r_{i,t}$ is daily individual futures contract returns over prior months *t-12* to *t-1* and $r_{m,t}$ is the corresponding daily return on the GSCI commodity futures contract market return index.
Skewness (*SKEWNEWS*)	*SKEWNESS* = skewness of daily returns over months *t-12* to *t-1*
Maximum daily return (*MAX*)	*MAX* = maximum daily futures contract returns over prior months *t-12 to t-1*
Expected shortfall (*ES*)	Individual futures contract returns expected shortfall measure at the 5% level over the prior months *t-12* to *t-1*.
BASIS (*BASIS*)	Daily $BASIS_t$ = second-nearby futures contract daily price/first-nearby futures contract daily price – 1.0 = $F_t^{T_2} / F_t^{T_1} - 1$, where $F_t^{T_1}$ is the end-of-month price of the first-nearby futures contract and $F_t^{T_2}$ is the end-of-month price of the second-nearby futures contract.

Monthly basis equals average of daily basis in prior month *t-1*.

Hedging pressure (*HPHE*)

The average of weekly hedging pressure ratio over prior months *t-12* to *t-1*:

$$HPHE_{week} = \frac{LONG_{hedgers,week} - SHORT_{hedgers,week}}{LONG_{hedgers,week} + SHORT_{hedgers,week}}$$

where *i* denotes hedgers. This follows Equation (1) of Fan et al. (2020).

Speculative pressure (*HPSP*)

The average of weekly speculative pressure ratios over prior months *t-12* to *t-1*:

$$HPSP_{week} = \frac{LONG_{speculators,week} - SHORT_{speculators,week}}{LONG_{speculators,week} + SHORT_{speculators,week}}.$$

where *i* denotes speculators. This follows Equation (1) of Fan et al. (2020).

3-month basis momentum (*BASM3*)

3-month basis momentum BASM3 for month *t* is measured as

$$BASM3_t = \prod_{j=1}^{3}(1+R_{t-j}^{T_1}) - \prod_{j=1}^{3}(1+R_{t-j}^{T_2}),$$

where $R_{t-j}^{T_1} = F_{t-j}^{T_1} / {}_{t-j-1}^{T_1} - 1.0$, $R_t^{T_2} = F_{t-j}^{T_2} / {}_{t-j-1}^{T_2} - 1.0$, $F_t^{T_1}$ is the end-of-month price of the first-nearby futures contract, and $F_t^{T_2}$ is the end-of-month price of the second-nearby futures contract.

12-month basis momentum (*BASM12*)

12-month basis momentum BASM12 for month *t* is measured as

$$BASM12_t = \prod_{j=1}^{12}(1+R_{t-j}^{T_1}) - \prod_{j=1}^{12}(1+R_{t-j}^{T_2})$$

where $R_{t-j}^{T_1} = F_{t-j}^{T_1} / {}_{t-j-1}^{T_1} - 1.0$, $R_{t-j}^{T_2} = F_{t-j}^{T_2} / {}_{t-j-1}^{T_2} - 1.0$, $F_t^{T_1}$ is the end-of-month price of the first-nearby futures contract, and $F_t^{T_2}$ is the end-of-month price of the second-nearby futures contract.

Beta with respect to the return on commodity market index (β_{CMKT})	β_{CMKT} is the slope coefficient from the following regression: $r_{i,t} = \beta_0 + \beta_{cmkt} r_{cmkt,\,t} + \varepsilon_{i,t}$, where $r_{i,t}$ is daily individual futures contract returns over prior months *t-36* to *t-1* and $r_{cmkt,\,t}$ is the corresponding daily return on the commodity market index.
Beta with respect to the return on the inflation index (β_{INF})	β_{INF} is the slope coefficient from the following regression: $r_{i,t} = \beta_0 + \beta_{INF} r_{INF,\,t} + \varepsilon_{i,t}$, where $r_{i,t}$ is daily individual futures contract returns over prior months *t-36* to *t-1* and $r_{INF,\,t}$ is the corresponding daily change on the commodity market producer price index.
Beta with respect to the return on USD index (β_{USD})	β_{USD} is the slope coefficient from the following regression: $r_{i,t} = \beta_0 + \beta_{USD} r_{USD,\,t} + \varepsilon_{i,t}$, where $r_{i,t}$ is daily individual futures contract returns over prior months *t-36* to *t-1* and $r_{USD,\,t}$ is the corresponding daily change on the USD index.
Futures contract trading volume in US$ terms (*VOLM*)	Average of daily trading volume of individual futures contracts in USD terms over prior months *t-12* to *t-1*
Futures contract open interest in US$ terms (*OPNI*)	Average of daily open interest of individual futures contracts in USD terms over prior months *t-12* to *t-1*
Amihud liquidity measure (*ALIQ*)	Daily Amihud measure = absolute value of daily futures contract return/daily trading volume $\times$1,000,000 = $\lvert RET \rvert / VOLM \times$1,000,000, *ALIQ* = average of daily measures over prior months *t-12* to *t-1*

Appendix B: The Definitions of Hedging Pressure and Speculative Pressure

In constructing the *CHP* risk factor, we consider the following three measures of hedging pressure and speculative pressure from hedgers and speculators, respectively. The first definition follows Fernandez-Perez et al. (2018):

$$HPHE_t^1 = \frac{LONG_{hedgers,t}}{LONG_{hedgers,t} + SHORT_{hedgers,t}},$$

$$HPSP_t^1 = \frac{LONG_{speculators,t}}{LONG_{speculators,t} + SHORT_{speculators,t}},$$

where *t* refers to calendar month. *LONG* and *SHORT* denote long and short positions, respectively. The second definition follows Fan et al. (2020):

$$HPHE_t^2 = \frac{LONG_{hedgers,t} - SHORT_{hedgers,t}}{LONG_{hedgers,t} + SHORT_{hedgers,t}},$$

$$HPSP_t^2 = \frac{LONG_{speculators,t} - SHORT_{speculators,t}}{LONG_{speculators,t} + SHORT_{speculators,t}}.$$

The third definition follows Kang et al. (2020):

$$HPHE_t^3 = \frac{LONG_{hedgers,t} - SHORT_{hedgers,t}}{OPNI_{hedgers,t}},$$

$$HPSP_t^3 = \frac{LONG_{speculators,t} - SHORT_{speculators,t}}{OPNI_{speculators,t}},$$

where *OPNI* refers to open interest. As in these earlier studies, we use the average weekly positions from CFTC reports over the past 12 months to calculate hedging pressure and speculative pressure. The following table summarizes the pair-wise correlations among three versions of hedging pressure and speculative pressure, respectively, from pooled time-series and cross-sectional observations (after taking the average in the past 12 months from *t*-12 to *t*-1):

we construct two-way sorted *CHP* risk factor as in Fernandez-Perez et al. (2018). For example, the long position consists of futures contracts with highest $HPSP_t^1$ (top 70%) and lowest $HPHE_t^1$ (bottom 30%). The short position consists of futures contracts with lowest $HPSP_t^1$ (bottom 30%) and highest $HPHE_t^1$ (top 70%). Then we form equal-weighted portfolio returns for month t to construct the *CHP* factor. We repeat the same procedure for the other two definitions of *HPHE* and *HPSP*. The three versions of *CHP* risk factors generate similar results.

Appendix C: Bootstrapping Procedures

Testing the null hypothesis that alpha does not depend on characteristics, $\Gamma_\alpha = 0_{L\times 1}$.

(1) Estimate the following model,

$$r_{t+1} = Z_t\Gamma_\alpha + Z_t\Gamma_\beta f_{t+1} + \varepsilon_{t+1}. \quad \text{(B1)}$$

We obtain Γ_α, Γ_β, f_{t+1} and ε_{t+1}, $t=0,...,T-1$. The IPCA estimation of Equation (B1) requires $Z_t^{'}Z_t$ and $Z_t^{'}r_{t+1}$ only. It does not require the use of *N*-dimension stock return r_{t+1} directly.

(2) Calculate the following residuals, d_{t+1}, $t=0,\ldots,T-1$, by multiplying $Z_t^{'}$ with fitted residuals ε_{t+1} from Equation (B1):

$$d_{t+1} = Z_t^{'}\varepsilon_{t+1} = Z_t^{'}r_{t+1} - [(Z_t^{'}Z_t)\Gamma_\alpha + (Z_t^{'}Z_t)\Gamma_\beta f_{t+1}]. \quad \text{(B2)}$$

(3) Let *boot* denote the number of bootstraps performed, *boot* = 1,…,1000. Obtain the sample of characteristic-weighted (Z_t-weighted) portfolio returns x_{t+1}^{boot} ($Z_t^{'}r_{t+1}^{boot}$) in the *boot*th step as

$$x_{t+1}^{boot} = Z_t^{'}r_{t+1}^{boot} = (Z_t^{'}Z_t)\Gamma_\beta f_{t+1} + d_{t+1}^{boot}, \quad \text{(B3)}$$

where d_{t+1}^{boot} is the *boot*th random draw from d_{t+1}, $t=0,\ldots,T-1$. In actual implementation, d_{t+1}^{boot} is multiplied by a student-*t* random variable with unit variance and five degrees of freedom. It is important to note in Equation (B3), we set Γ_α = 0 as required under the null hypothesis. Equation (B3) indicates that x_{t+1}^{boot} is constructed using residuals from the unrestricted model in Equation (B1), but the fitted value of return $\hat{r}_{t+1}$ (i.e., bootstrapped return r_{t+1}^{boot}) in Equation (B3) excludes $Z_t\Gamma_\alpha$ because we now set Γ_α = 0. Therefore, characteristic-weighted returns $Z_t^{'}r_{t+1}^{boot}$ drop out the impact of $Z_t\Gamma_\alpha$. This is the null hypothesis that $\Gamma_\alpha = 0_{L\times 1}$. We bootstrap characteristic-weighted returns under the null that characteristics do not impact alpha.

(4) Use the bootstrapped sample of characteristic-weighted portfolio returns x_{t+1}^{boot} in the *boot*th step to estimate the unrestricted model in Equation (B1) again and obtain Γ_α^{boot}. The estimation only requires characteristic-weighted returns x_{t+1}^{boot}. The estimation does not require *N*-dimension stock returns r_{t+1}^{boot}.

(5) Calculate the Wald-like statistic from the *boot*th step bootstrap as $W_\alpha^{boot} = (\Gamma_\alpha^{boot})'(\Gamma_\alpha^{boot})$. The original Wald-like statistic from the unrestricted model in Equation (B1) is $W_\alpha = (\Gamma_\alpha)'(\Gamma_\alpha)$.

(6) Calculate the *p*-value for testing the null hypothesis that $\Gamma_\alpha = 0_{L\times 1}$ as the percentage of W_α^{boot} statistics that are larger than W_α statistics.

Testing the null hypothesis that the l^{th} characteristic has no impact on the betas of all K latent factors, $\gamma_{\beta,l} = 0_{1\times K}$.

(1) Estimate the following restricted model after we cannot reject the null hypothesis that $\Gamma_\alpha = 0_{L\times 1}$,

$$r_{t+1} = Z_t \Gamma_\beta f_{t+1} + \varepsilon_{t+1}. \tag{B4}$$

(2) Calculate the following residuals, d_{t+1}, $t=0,\dots,T-1$, by multiplying $Z_t^{'}$ with fitted residuals ε_{t+1} from Equation (B4):

$$d_{t+1} = Z_t^{'} \varepsilon_{t+1} = Z_t^{'} r_{t+1} - (Z_t^{'} Z_t)\Gamma_\beta f_{t+1}. \tag{B5}$$

(3) Let *boot* denote the number of bootstraps performed, *boot* = 1,…,1000. Obtain the sample of characteristic-weighted (Z_t-weighted) portfolio returns x_{t+1}^{boot} ($Z_t^{'} r_{t+1}^{boot}$) in the *boot*th step as

$$x_{t+1}^{boot} = Z_t^{'} r_{t+1}^{boot} = (Z_t^{'} Z_t)\Gamma_\beta^{reset} f_{t+1} + d_{t+1}^{boot}, \tag{B6}$$

where d_{t+1}^{boot} is the *boot*$^{\text{th}}$ random draw from d_{t+1}, $t=0,\dots,T-1$. In actual implementation, d_{t+1}^{boot} is multiplied by a student-t random variable with unit variance and five degrees of freedom.

Notice here the l^{th} row of Γ_β^{reset} is set to be $0_{1\times K}$. Let $\gamma_{\beta,l}^{reset}$ be the l^{th} row of Γ_β^{reset}. We impose the condition that $\gamma_{\beta,l}^{reset} = 0_{1\times K}$ under the null hypothesis. This implies that the l^{th} characteristic has no impact on the beta of all latent factors. Then we estimate the model in Equation (B4) again using bootstrapped x_{t+1}^{boot} to obtain Γ_β^{boot}, where $\gamma_{\beta,l}^{boot}$ is the l^{th} row of Γ_β^{boot}

(4) Calculate the Wald-like statistic from each bootstrap as $W_{\beta,l}^{boot} = (\gamma_{\beta,l}^{boot})'(\gamma_{\beta,l}^{boot})$. The Wald-like statistic using the actual return data from the model in Equation (B4) is calculated as $W_{\beta,l} = \gamma_{\beta,l}^{'} \gamma_{\beta,l}$.

(5) Calculate the p-value for testing the null hypothesis that $\gamma_{\beta,l} = 0_{1\times K}$ as the percentage of $W_{\beta,l}^{boot}$ statistics that are larger than $W_{\beta,l}$ statistics.

Testing the null hypothesis that the m^{th} observable risk factor has no incremental explanatory power, $\gamma_{\delta,m} = 0_{M\times 1}$.

(1) Estimate the following restricted model after we cannot reject the null hypothesis that $\Gamma_\alpha = 0_{L\times 1}$,

$$r_{t+1} = Z_t \Gamma_\beta f_{t+1} + Z_t \Gamma_\delta g_{t+1} + \varepsilon_{t+1}. \qquad \text{(B7)}$$

(2) Calculate the following residuals d_{t+1}, $t=0,\ldots,T-1$, by multiplying $Z_t^{'}$ with fitted residuals ε_{t+1} from Equation (B7):

$$d_{t+1} = Z_t^{'} \varepsilon_{t+1} = Z_t^{'} r_{t+1} - (Z_t^{'} Z_t) \Gamma_\beta f_{t+1} - (Z_t^{'} Z_t) \Gamma_\delta g_{t+1}. \qquad \text{(B8)}$$

(3) Let *boot* denote the number of bootstraps performed, *boot* = 1,…,1000. Obtain the sample of characteristic-weighted (Z_t-weighted) portfolio returns x_{t+1}^{boot} ($Z_t^{'} r_{t+1}^{boot}$) in the $boot^{th}$ step as

$$x_{t+1}^{boot} = Z_t^{'} r_{t+1}^{boot} = (Z_t^{'} Z_t) \Gamma_\beta f_{t+1} + (Z_t^{'} Z_t) \Gamma_\delta^{reset} g_{t+1} + d_{t+1}^{boot}, \qquad \text{(B9)}$$

where d_{t+1}^{boot} is the $boot^{th}$ random draw from d_{t+1}, $t=0,\ldots,T-1$. In actual implementation, d_{t+1}^{boot} is multiplied by a student-*t* random variable with unit variance and five degrees of freedom.

Note here the m^{th} column of Γ_δ^{reset} is set to be $0_{M\times 1}$. Let $\gamma_{\delta,m}^{reset}$ be the m^{th} column of Γ_δ^{reset}. We impose the condition that $\gamma_{\delta,m}^{reset} = 0_{M\times 1}$ under the null hypothesis. This implies that the m^{th} observable risk factor has no incremental explanatory power. Then we estimate the model in Equation (B7) again using bootstrapped x_{t+1}^{boot} to obtain Γ_δ^{boot}, where $\gamma_{\delta,m}^{boot}$ is the m^{th} column of Γ_δ^{boot}.

(4) Calculate the Wald-like statistic from each bootstrap as $W_{\delta,m}^{boot} = (\gamma_{\delta,m}^{boot})'(\gamma_{\delta,m}^{boot})$. The Wald-like statistic using the actual return data from the model in Equation (B7) is calculated as $W_{\delta,m} = \gamma_{\delta,m}^{'} \gamma_{\delta,m}$.

(5) Calculate the *p*-value for testing the null hypothesis that $\gamma_{\delta,m} = 0_{M\times 1}$ as the percentage of $W_{\delta,m}^{boot}$ statistics that are larger than $W_{\delta,m}$ statistics.

Appendix D: Panel Regressions of Futures Contract Monthly Returns and Market Betas $BETA_{CMKT}$ on Futures Contract Characteristics

The sample covers the period from January 1981 to June 2022. Panel A reports the panel regressions of monthly returns on individual commodity futures contracts on different combinations of contract characteristics. The benchmark model in the first row includes three characteristics, *MOM12*, *ES*, and *IVOL*. Then the panel regressions add in each of the remaining 17 characteristics, one at a time. The 20 commodity futures contract characteristics include *MOM3*, *MOM12*, *CTR18*, *CTR36*, *R52WH*, *IVOL*, *SKEWNESS*, *MAX*, *ES*, *BASIS*, *HPHE*, *HPSP*, *BASM3*, *BASM12*, $BETA_{CMKT}$, $BETA_{INF}$, $BETA_{USD}$, *VOLM*, *OPNI*, and *ALIQ*. All characteristics are measured prior to month *t* when equally weighted portfolio returns are constructed. The details of the construction of the sorting variables are provided in Appendix A. Panel B reports the panel regressions of estimated market betas ($BETA_{CMKT}$) on individual commodity futures contracts on different combinations of contract characteristics. Market betas are estimated from months *t*+1 to *t*+12. $BETA_{CMKT}$ is estimated with respect to returns on the commodity market index. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A: The Dependent Variable is Monthly Returns of Individual Futures Contracts from *t*+1

First 3 Characteristics			*4th Characteristic*		R^2
MOM12	*ES*	*IVOL*			
0.010***	-0.005	-0.010*			0.002
0.008***	-0.005	-0.010*	*MOM3*	0.004	0.002
0.019***	-0.005	-0.011***	*CTR18*	-0.012***	0.003
0.014***	-0.005	-0.010*	*CTR36*	-0.008***	0.003
0.005	-0.007	-0.010*	*R52WH*	0.007*	0.002
0.009***	-0.002	-0.007	*SKEWNESS*	-0.009***	0.003
0.010***	-0.012***	-0.008	*MAX*	-0.010***	0.003
0.005	-0.003	-0.010	*HPHE*	-0.008***	0.002
0.005	-0.003	-0.009	*HPSP*	0.005	0.002
0.009***	-0.004	-0.007	*BASIS*	-0.002	0.002
0.010***	-0.003	-0.007	*BASM3*	-0.007	0.002
0.009***	-0.004	-0.007	*BASM12*	0.003	0.002
0.010***	-0.008	-0.012***	$BETA_{CMKT}$	-0.003	0.002
0.009***	-0.007	-0.012***	$BETA_{INF}$	-0.005*	0.002
0.009***	-0.004	-0.009*	$BETA_{USD}$	-0.000	0.002
0.008***	-0.005	-0.011*	*VOLM*	-0.008***	0.002
0.008***	-0.004	-0.009*	*OPNI*	-0.007***	0.002
0.008***	-0.004	-0.011*	*ALIQ*	0.007***	0.002

Panel B: The Dependent Variable is Estimated Betas from months *t*+1 to *t*+12

First 3 Characteristics			*4th Characteristic*		R^2
MOM12	*ES*	*IVOL*			
0.220***	-1.510***	-1.337***			0.262
0.199***	-1.510***	-1.336***	*MOM3*	0.044	0.262
0.179***	-1.508***	-1.327***	*CTR18*	0.059	0.262
0.216***	-1.503***	-1.309***	*CTR36*	0.051	0.264
0.249***	-1.499***	-1.337***	*R52WH*	-0.042	0.262
0.232***	-1.577***	-1.388***	*SKEWNESS*	0.160***	0.266
0.210***	-1.336***	-1.380***	*MAX*	0.263***	0.267
0.326***	-1.529***	-1.289***	*HPHE*	0.133	0.279
0.343***	-1.531***	-1.306***	*HPSP*	-0.145	0.279
0.236***	-1.439***	-1.261***	*BASIS*	-0.009	0.236
0.238***	-1.440***	-1.261***	*BASM3*	0.007	0.235
0.246***	-1.440***	-1.262***	*BASM12*	-0.040	0.234
0.192***	-1.074***	-1.021***	$BETA_{CMKT}$	0.546***	0.299
0.232***	-1.372***	-1.189***	$BETA_{INF}$	0.213***	0.269
0.220***	-1.451***	-1.276***	$BETA_{USD}$	-0.073	0.262
0.242***	-1.348***	-1.195***	*VOLM*	0.223***	0.245
0.242***	-1.382***	-1.237***	*OPNI*	0.161	0.239
0.244***	-1.380***	-1.205***	*ALIQ*	-0.182***	0.242