Benchmarking Commodities

JESSE BLOCHER,* RICKY COOPER,** AND MARAT MOLYBOGA***

ABSTRACT

While much has been said about the financialization of commodities, much less is known about how to profitably invest in commodities. Existing studies of Commodity Trading Advisors (CTAs) do not adequately address this question because only 19% of CTAs invest solely in commodities, despite their name. We develop a four-factor asset pricing model to benchmark commodity investment, and compare it to the existing factor model benchmarks currently used to evaluate CTAs. Only our four-factor model prices both commodity spot and term risk premia, as measured by a GRS test and R^2. Overall, our four-factor model prices commodity risk premia better than the popular Fama-French three-factor model prices equity risk premia, and thus is an appropriate benchmark with which to evaluate commodity investment vehicles.

Current version: February 22, 2016

*Assistant Professor of Finance, Owen Graduate School of Management, Vanderbilt University, 401 21st Avenue South, Nashville, TN 37203. **Assistant Professor of Finance, Stuart School of Business, Illinois Institute of Technology. ***Chief Risk Officer and Director of Research, Efficient Capital Management and Adjunct Faculty at the Stuart School of Business, Illinois Institute of Technology. Blocher acknowledges support from the Chicago Mercantile Exchange and Vanderbilt's Financial Markets Research Center. We are grateful for comments and suggestions from Craig Lewis, Nick Bollen, Bob Whaley, Luke Taylor, Adam Reed, Clemens Sialm, Vikas Agarwal, Jeff Busse, and Sohpie Shive. We benefitted from conversations with Christophe L'Ahelec. Cheng Jiang provided excellent research assistance. The authors can be contacted via email at: jesse.blocher@owen.vanderbilt.edu, rcooper3@stuart.iit.edu, and molyboga@efficientcapital.com.

Benchmarking Commodities

> *Like most products in the liquid alternatives space, this is not a simple plug-and-play category, where any above-average fund will suffice … The challenge, as always, is finding the right manager. But it doesn't help that the managed futures space is still a very long way from enabling simple and straightforward comparisons.* - Investment News, Jan 14, 2015, "Managed futures funds shine anew, but mystery remains"

> *Commodities "couldn't be hated more"…Four years of negative returns for indices tracking futures, with a fifth under way, have undermined the idea that leaving part of one's portfolio in a basket of oil, natural gas, soyabeans, copper and other commodities was prudent. "There's zero interest right now from the institutional space," says Lawrence Loughlin of Drobny Capital*. – Financial Times, June 3, 2015, "Investment: Revaluing Commodities"

The literature on commodities dates back at least to Keynes (1923), but most of it focuses on production and storage decisions or the role of commodities in international trade (Rouwenhorst and Tang (2012)). There is a large and growing literature around the financialization of commodities, the purported cause of which is increasing investment by finance professionals or so-called 'speculators' (e.g. Cheng and Xiong (2014)). However, there has been less research about how astute investors should incorporate commodities into a diversified portfolio. Since the global capital (institutional and retail) allocated to commodities is approximately $330B, this is an important question.[1]

Gorton and Rouwenhorst (2006) and Erb and Harvey (2006) advocated for an equal-weighted index of commodity futures, an easily-implemented passive strategy. But this approach

[1] $330B comes from investment report from Barclays Capital Commodities Research via a HewittEnnisKnupp Global Invested Capital Report, June 2014.

has yielded negative or zero returns over much of its history, and practitioners are abandoning it.[2] Figure 1 shows this poor performance of an equally weighted market index since 1987.

Some papers have examined Commodity Trading Advisors (CTAs) (e.g. Fung and Hsieh (1997), Fung and Hsieh (2000), Bhardwaj, Gorton, and Rouwenhorst (2014)), but have done so using diversified factor models because only 19% of CTAs invest exclusively in commodities, despite their name (see Table I). Fittingly, these studies have typically been interpreted as research on hedge funds, rather than commodities (e.g., Kosowski, Naik, and Teo (2007), Bollen and Whaley (2009)). Narrowing the portfolio decision to this 19% of CTAs that invest solely in commodities (or "commodity funds") may be the best way to incorporate commodities into a diverse portfolio, but no commodity-specific benchmark exists with which to evaluate these managers. Since Roll (1978) showed that different benchmarks can yield different rankings of 'skill', independently identifying the right benchmark is a necessary first step to manager selection.

This paper establishes a parsimonious, tradeable, four-factor model benchmark, with which investors can evaluate commodity fund managers (or other commodity investments, such as Exchange Traded Funds, or ETFs). Our model not only prices commodity spot risk premia, but also commodity term risk premia, identified by Szymanowska et al. (2014). Our four-factor model fails to price only two test assets among five different four-way portfolio sorts (two spot premia, three term premia, for a total of twenty portfolios). This performance exceeds that of the Fama-French 3 factor model (Fama and French (1993)) and is on par with that of the new Fama-French 5 factor model (Fama and French (2015)), both in the equity space.

[2] Investment: Revaluing Commodities, June 3, 2015, Financial Times. http://www.ft.com/cms/s/0/a6ff2818-094c-11e5-8534-00144feabdc0.html. Also see Bhardwaj, Gorton, and Rouwenhorst (2015).

The four factors in our model include a market factor, a time series momentum factor, and separate high and low term premia factors, sorted on commodity basis. These factors are drawn from the extant literature and based in commodity fundamentals, and each has been shown separately to capture a risk premium embedded in commodity futures, though never together in the form we propose.

To establish our four-factor model as necessary, we run a horserace between our model and two popular models established in the literature. The first is the popular model of Fung and Hsieh (Fung and Hsieh (2001)), which we call the FH model. This model covers a wide variety of strategies and is intended as a descriptive model to identify the strategies used by hedge funds and CTAs. One drawback of this model is that the factors are not tradeable, making interpretation difficult. This model has also been criticized by Bhardwaj, Gorton, and Rouwenhorst (2014), who argue that the negative performance of the factors means alpha identified based on this model is spurious.[3] The second model comes from Bhardwaj, Gorton, and Rouwenhorst (2014), who include factors for commodities, interest rate derivatives, and currency futures. Since our focus is on commodities, we only test the model's commodity factors, which we call the BGR model.[4]

We find that both our model and the BGR model price spot risk premia adequately. Both estimate an alpha equal to zero for all test assets, fail to reject the GRS test that all portfolio alphas are jointly set to zero Gibbons, Ross, and Shanken (1989), and have high adjusted R^2. The FH model fails to price several of the spot premia test assets and the GRS test rejects null

[3] In some factor models, negative factor loadings could give positive performance on a poorly performing factors. But since the Fung-Hseih factors are not tradeable, this interpretation is not applicable.

[4] Bhardwaj, Gorton, and Rouwenhorst (2014) do not test their model, but simply assert it as capturing known trading patterns in commodities. The commodity-only version of their model is also almost identical to a model proposed and tested more thoroughly in a working paper by Bakshi, Gao Bakshi, and Rossi (2014), thus we can refer to it as the "BGR" model and use that to refer to both papers.

hypothesis of joint zero alpha for all portfolio sorts. The adjusted R^2 for the FH model is zero for all test portfolios.

Our four-factor model is the only model that can price term premia. The BGR model can only price 2 of 12 test asset portfolios, and the FH model can price 3 of 12. In contrast, our four-factor model including the two term premia factors successfully price 10 of 12 test asset portfolios. At both the four -and six-month horizon, a GRS test of our four-factor model fails to reject the null of zero alpha for all portfolios. At a two-month horizon, our four-factor model prices 3 of 4 portfolios.

Until now, benchmarking commodity investments has been inhibited by a lack of understanding of the drivers of risk premia. Recently, however, the literature has coalesced around a few key drivers of commodity risk premia, represented by the four factors in our model. Simultaneously, increased interest in commodity investment in the past decade combined with the poor performance of passive market indexes means sophisticated investors are more interested in evaluating the performance of active commodity fund managers. Financial advisors have even suggested that individuals include commodities in their personal asset allocation.[5] Yet, to our knowledge, there is not a thoroughly tested and established benchmark to evaluate commodity fund managers or commodity ETFs.

While commodity investment often is included as a subset of the hedge fund/CTA literature, there are many similarities between commodity markets and equity markets. Both are publicly traded and marked-to-market daily (in most cases). Both have a clearly identified regulatory body (the CFTC and SEC). Both represent a defined investment set within which a manager (Commodity Fund or Mutual Fund) must choose either long or short positions. Given

[5] "Speculating on commodities can add diversity to your portfolio," *The Financial Times*, June 16, 2015. http://www.ft.com/intl/cms/s/2/eee82070-ea99-11e4-96ec-00144feab7de.html

these similarities, our paper can be seen as establishing a factor model benchmark for Commodity Funds in the same way that Fama and French (1992, 2015) have established a benchmark for Mutual Funds.

Our paper proceeds as follows. In section I, we motivate and describe our factor model, including its basis in the current literature. In section II, we establish the test criteria and test portfolios we use to compare asset pricing models. In section III, we discuss the results of those tests and section IV concludes.

I. A factor model of commodity returns

In this first section, we select factors already established in the literature and adjust them for use in benchmarking monthly commodity fund returns. Therefor, while each of our factors has some precursor in the literature, to our knowledge, nobody has combined them together into a single, parsimonious, tradeable benchmark and tested the combined performance at pricing commodity futures spot and term premia.

We follow the general contribution of Szymanowska et al. (2014) and use two term premium factors to account for the futures basis. However, we show that two summarized term premium factors are sufficient instead of the six proposed in that paper. Furthermore, we show that a spot basis factor is redundant. We then add a market factor, which is an equal-weighted portfolio of all commodity futures and is common in the commodities literature (e.g., Gorton and Rouwenhorst (2006), Erb and Harvey (2006)). We also include a time series momentum factor, which is also present in several commodities papers (e.g., Moskowitz, Ooi, and Pedersen (2012), Miffre and Rallis (2007)), but is in contrast with cross-sectional momentum common in equities (Jegadeesh and Titman (2011)) and in some commodity models (e.g. the BGR model in both Bakshi, Gao Bakshi, and Rossi (2014) and Bhardwaj, Gorton, and Rouwenhorst (2014)). We discuss the differences in these two momentum factors in detail later in this section. Finally, we use a monthly time series of commodity futures returns, in contrast to Szymanowska et al. (2014), who use bimonthly returns and holding periods of up to eight months. Overall, these adjustments result in fewer factors that are easier to implement, tradeable, and have comparable explanatory power. We next discuss our data before describing these factors in more detail.

A. Commodity risk premia and motivation for factor selection

Explaining commodity risk premia dates back at least to Keynes (1923), who proposed a theory of "normal backwardation," in which short hedgers of commodities outnumber long hedgers such that natural hedgers are net short. Thus, the assumed natural state of the market is for futures prices to be lower than expected future spot prices to give speculators a positive expected return for assuming the price risk.[6] In this common "insurance" view of commodity futures risk premia, commodity futures traders accept price risk from hedgers in exchange for a risk premium. Rouwenhorst and Tang (2012) survey the extensive literature and conclude that evidence for this theory is weak.

Keynes' theory predates modern asset pricing theory, embodied in the capital asset pricing model. Early studies find little evidence that this model applies to commodities markets (e.g., Dusak (1973), Carter, Rausser, and Schmitz (1983)), and recent studies by Gorton and Rouwenhorst (2006), Rouwenhorst and Tang (2012), Erb and Harvey (2006) confirm these findings. The explanation of commodity risk premia in the context of the capital asset pricing

[6] Note that normal backwardation (futures price < expected future spot price) differs from backwardation (futures price < current spot price).

model remains an open question in the commodities literature, and the literature moved to arbitrage pricing models.[7]

The strongest empirical evidence around commodity risk premia associates inventory with commodity risk premia in the context of the theory of storage, which dates back to Kaldor (1939), Working (1949), and Brennan (1958). This theory links commodities futures prices to the storage decisions of inventory holders, in terms of financing and warehousing costs net a convenience yield. Gorton, Hayashi, and Rouwenhorst (2012) investigate the fundamentals of commodity investing and find that inventory and storage are the key fundamentals in pricing commodity risk premia and that both correlate with the commodity future's basis. This finding features prominently in our benchmark four-factor model.

Finally, and most recently, Szymanowska et al. (2014) investigate the term structure of commodity risk premia and show the existence of term premia in commodity futures. They show that factors derived from a sort on futures basis can explain these premia, but use two factors per maturity at two, four, and six months for a total of six factors. We interpret these factors as capturing expected future changes in commodity inventories given the link between basis and inventory levels established by Gorton, Hayashi, and Rouwenhorst (2012).

B. Data and computation of futures premia and returns

We use 21 different commodity futures from Commodity Systems Inc. that represent all major sub-sectors of commodity markets (i.e., energy, agricultural, and metals). The contracts include Soybean Oil, Corn, Cocoa, Light Crude Oil, Cotton, Gold, Copper, NY Harbor ULSD (Heating Oil), Coffee, Lumber, Hogs, Oats, Orange Juice, Soy Beans, Silver, Soy Meal, Wheat

[7] Some literature relates forwards and futures premia to the consumption capital asset pricing model (for example, see Cooper (1993)), which reports that forward and futures contracts respond to time-varying risk premium formulations.

(CBT only), Feeder Cattle, Live Cattle, Gasoline RBOB, and Rice Rough for the period between September 1987 and December 2014. Table II provides information about Bloomberg codes and exchanges associated with each futures market.

In constructing our factors, we follow convention and consider the spot price to be the price of the contract nearest to expiration and expiring at least two months from the current month. This avoids liquidity problems, which can plague the pricing of shorter maturity contracts. The two-month, four-month, and six-month contracts are then defined as the first contract to expire at least two months, four months, and six months after the spot contract expires.[8] From the commodity price series, we construct several variables from which all the model's factors are constructed. We define the spot premium of the commodity as the change in the logarithm of the spot price, $s_i(t)$. Therefore, the realized spot premium of commodity *i* at time *t*, $\hat{\pi}_{s,i}(t)$, is defined as

$$\hat{\rho}_{s,i}(t) = \ln\left[s_i(t)\right] - \ln\left[s_i(t-1)\right]. \tag{1}$$

As is standard in the literature, this premium formulation excludes all returns on the required collateral. Intuitively, these returns are comparable to returns in excess of the risk-free rate, because collateral is typically reinvested at that rate.

The *n*-month basis for commodity *i*, $y_i^n(t)$ is defined as the logarithm of the ratio of the *n*-month futures price $f_i^n(t)$ to the spot price. Generally, the *n*-month maturity term premium $\hat{\pi}^n_{y,i}(t)$ is defined as the change in this value:

$$\hat{\rho}^n_{y,i}(t) = y_i^n(t) - y_i^{n*}(t-1) = \ln[f_i^n(t) \big/ f_i^{n*}(t-1)] - \ln[s_i(t)/s_i(t-1)] \tag{2}$$

[8] As a simple example, corn has contracts expiring in months 3, 5, 7, 9, and 12. In October (10), the spot contract will be December (12), the 2-month contract will be March (3), the 4-month contract May (5), and the 6-month contract July (7). Some commodities also have monthly expirations, in which case some expiration months would be skipped on any given date. This approach is similar to Szymanowska et al. (2014).

This may be thought of as a calendar spread, which is computed by buying the *n*-month futures contract and shorting the spot futures contract. The futures returns themselves may be written as

$$r_{f,i}^{n}(t) = \ln\left[f_i^{n-1}(t)\right] - \ln\left[f_i^{n}(t-1)\right].^{9} \tag{3}$$

The cost-of-carry relationship for the futures markets allows us to break the *n*-month expected futures return for commodity *i* into a spot premium and a term premium.[10] The cost of carry model may be defined as

$$f_i^{n}(t) = s_i(t)e^{\int_t^{t+n} y_i(t)dt} \tag{4}$$

where y(t), the time *t* instantaneous cost of carry includes the risk-free rate, the storage rate for the commodity *i*, and a generally negative rate known as the convenience yield. The spot price $s_i(t)$ is the true underlying commodity price, and $f_i^{n}(t)$ is the futures price with maturity *n*. The total cost of carry over the life of the contract is summarized by the basis, defined as

$$y_i^{n} = \ln\left[f_i^{n}(t)\right] - \ln\left[s_i(t)\right] = \int_t^{t+n} y_i(t)\,dt\ . \tag{5}$$

Taking derivatives and rearranging yields the equation

$$d\ln\left[f_i^{n}(t)\right] = d\ln\left[s_i(t)\right] + y_i^{n}\ . \tag{6}$$

If we now consider small discrete time changes (so that equation 6 is still approximately correct), then we can write the expected spot premium as

$$\pi_{s,i}(t) = E_t\left[\ln(s_i(t+1) - \ln(s_i(t) - y_i^{1}(t)\right], \tag{7}$$

and the expected term premium as

[9] *n* does not always decrement to *n-1*, because contracts do not necessarily expire every month.

[10] Erb and Harvey (2006), Routledge, Seppi, and Spatt (2000), and Fama and French (1987) establish a link between basis and commodity futures risk premia.

$$\pi_{y,i}^{n}(t) = E_t[y_i^{n-1}(t) + y_i^1] - y_i^n \quad (8)$$

Equation 7 gives the premium as the difference between the expected change in the spot price and the one-period basis. Equation 8 gives the premium as the deviation from the expectation hypothesis.

We can now expand equation 6 to expected futures return as

$$E_t[r_{f,i}^n(t+1)] = E_t[f_i^{n-1}(t+1) - f_i^n(t)] = E[s_i(t+1) + y_i^{n-1}(t+1) - s_i(t) - y_i^n(t) + y_i^1(t) - y_i^1(t)] \quad .(9)$$

This reduces into

$$E_t[r_{f,i}^n(t+1)] = \pi_{s,i}(t+1) + \pi_{y,i}^n(t+1), \quad (10)$$

which is in terms of risk premia. Recall that we define the spot commodity as the nearest term futures contract, and this definition includes the true spot price plus the one-period cost of carry. Thus, our spot premium and term premia measures correspond to realizations of the premia in equation 10.

C. Factor selection and construction – spot premia

We consider factors for each premium in turn, starting with the spot premium. We first include a market factor (MKT), which is an equally weighted average of all commodities' one period spot return. Our market factor is

$$MKT(t) = \frac{1}{N}\sum_{i=1}^{N}\hat{\pi}_{s,i}(t) \quad (12)$$

where N, the number of total commodities, is 21. This follows the well-known results of Erb and Harvey (2006) and Gorton and Rouwenhorst (2006) who show the value of an equal-weighted market index.

Next, we include a momentum factor, shown necessary by Gorton, Hayashi, and Rouwenhorst (2012), among others.[11] We compute a time series momentum factor (TSMOM) as in Moskowitz, Ooi, and Pedersen (2012), which is the difference in return between an equally weighted portfolio of commodities with a positive return over the previous twelve months and one with a negative return over the previous twelve months. Specifically, we define momentum as

$$TSMOM(t) = \left[\frac{1}{N_{pos}} \sum_{i \in P} \left[\hat{\pi}_{s,i}(t) \right] - \frac{1}{N_{neg}} \sum_{i \in L} \left[\hat{\pi}_{s,i}(t) \right] \right] \quad (13)$$

where *neg* and *pos* refer to the set of commodities with positive and negative trailing 12-month returns. N_{pos} and N_{neg} refer to the number of commodities in each respective group.
Time series momentum differs from typical momentum measures (sometimes called cross-sectional momentum) in the selection of the high and low momentum portfolios. Time series momentum sorts based on sign: only positive momentum stocks are placed in the high momentum portfolio; only negative momentum stocks are placed in the low momentum portfolio. Cross-sectional momentum uses a ranking, taking those in the top decile as the high momentum portfolio and the bottom decile as low momentum. The MOM factor is computed as

$$MOM(t) = \frac{1}{N_g} \left[\sum_{i \in H} \hat{\pi}_{s,i}(t) - \sum_{i \in L} \hat{\pi}_{s,i}(t) \right], \quad (15)$$

where H is the set of High group commodities when sorted on past 12 month return, and L is the set of Low group commodities, and N_g is equal to the number of commodities in each group. The two groups in MOM have the same size by definition, in contrast to TSMOM, which allows differently sized groups.

[11] Others who find evidence for some type of momentum factor include Erb and Harvey (2006), Asness, Moskowitz, and Pedersen (2013), Fuertes, Miffre, and Rallis (2010), and Miffre and Rallis (2007).

Conceptually, cross-sectional momentum is capturing relative momentum in the cross section, whereas time series momentum captures each security's own trend. In equity markets, the differences are not material, because the large pool of securities ensures that a High cross-sectional momentum portfolio is constructed almost entirely of stocks with positive momentum, and the Low cross-sectional momentum portfolio is constructed almost entirely of stocks with negative momentum. However, in commodity, foreign exchange and equity index futures markets, where there are significantly fewer securities (e.g. 21 in our sample), time series momentum can differ greatly from cross-sectional momentum. This happens when, for instance, there are only one or two futures contracts with either positive or negative momentum. In those cases, the Low cross-sectional momentum portfolio could contain contracts with zero or positive momentum, and the High cross-sectional momentum portfolio could contain contracts with zero or negative momentum. This situation occurs because the sorting does not filter on sign, it simply ranks from high to low.

We confirm that time-series momentum is a much better predictor of commodity risk premia than cross-sectional momentum, in line with the literature (e.g., Moskowitz, Ooi, and Pedersen (2012), Miffre and Rallis (2007), Baltas and Kosowski (2012)). This choice of the momentum factor is one of the differences between our model and that of Bakshi, Gao Bakshi, and Rossi (2014) and Bhardwaj, Gorton, and Rouwenhorst (2014). While we choose the time series momentum factor, they both use cross-sectional momentum.

Szymanowska et al. (2014) motivate and derive a high-minus-low factor to explain spot premia. Specifically, the HML factor is defined as

$$HML(t) = \frac{1}{N_g}\left[\sum_{i\in H}\hat{\pi}_{s,i}(t) - \sum_{i\in L}\hat{\pi}_{s,i}(t)\right] \tag{11}$$

where H is the set of commodities with a spot return in the High group, L is the set of commodities with a spot return in the Low group, and N_g is the number of commodities in each group. Bhardwaj, Gorton, and Rouwenhorst (2014) split the commodities at the median, and since we have 21 total commodities, N_g is 10. Bakshi, Gao Bakshi, and Rossi (2014) derives a similar factor but set N_g equal to 5. The HML factor we estimate follows the former definition using the median.

Overall, we consider four possible spot premia factors: MKT, TSMOM, MOM, and HML. We conclude, ultimately, to only include MKT and TSMOM in addition to the two term premia factors discussed in the next section. We present the results justifying this choice later with asset pricing tests. Rouwenhorst and Tang (2012) summarize a long and growing literature of commodities factors that price commodity returns. We do not comprehensively test all possibilities, but rather focus on the factors that most frequently show up as important in pricing commodity risk premia.

D. Factor selection and construction – term premia

We next consider the term premium. To price the term premium, we choose two factors. First, we construct a high-term premium factor (H_{term}) consisting of the average of the 2-month, 4-month, and 6-month realized term premia for the 10 commodities with above-median basis (as previously defined in the HML factor). We also construct a low-term premium factor (L_{term}), computed the same way as H_{term}, except using the 10 commodities with below-median basis. These two factors are defined as

$$
\begin{aligned}
H_{term}(t) &= \frac{1}{N_g}\sum_{i\in H}\left[\frac{1}{3}\sum_{n=2,4,6}\hat{\pi}^{n}_{y,i}(t)\right] \\
L_{term}(t) &= \frac{1}{N_g}\sum_{i\in L}\left[\frac{1}{3}\sum_{n=2,4,6}\hat{\pi}^{n}_{y,i}(t)\right],
\end{aligned}
\tag{14}
$$

where H is the set of commodities with above-median basis, and L is the set of commodities with below-median basis. N_g is the number of commodities in each group, which is 10.

These two factors follow the intuition of Szymanowska et al. (2014), who also construct their longer term basis factor as separate high and low factors to explain commodity term premia. However, Szymanowska et al. (2014) compute their term structure basis factor using so-called “spreading” returns that span the maturity difference of the computed term premia. Thus, to explain 2-, 4-, and 6-month term premia, they require three H factors and three L factors, each with maturities matching those three holding periods. This approach has two implications. First, it is not obvious that a set of factors designed to explain multi-month holding period returns will adequately explain 1-month returns when applied to CTAs or ETFs. Second, including six additional factors in a single benchmark model for commodity funds is unwieldy and likely redundant. Our goal is to preserve the economic intuition and econometric relationships while creating easily implementable factors. Our results demonstrate that our factors, though simpler, maintain power in explaining futures returns. Intuitively, our term premium factors capture the equally weighted average of the *expected change in spot basis* across different time horizons. Because of the link between basis and inventories, this naturally maps to trader information about the expected evolution of commodity inventories, which is a key fundamental driver of commodity risk premia.

Table III summarizes our factors as well as the factors in our competing models (the FH model and BGR model, discussed later). For all asset pricing tests, we apply Newey and West (1987) corrections for heteroskedasticity and autocorrelation with 12 lags, because there is a pronounced seasonal effect in commodities (Gorton, Hayashi, and Rouwenhorst (2012)). The monthly excess return for almost every factor is modestly positive, but statistically different from zero. Only the market factor and Fung-Hsieh Commodity Factor (FHCOM) are negative, but both are insignificant.

Additionally, Figure 1 shows a time series plot of the performance of our factors. Bhardwaj, Gorton, and Rouwenhorst (2014) critique Fung and Hsieh (2001) because the poor performance of their factors can spuriously indicate a fund delivers alpha. This is because Fung and Hsieh's factors are not tradeable and therefore cannot be shorted (i.e. cannot have negative coefficients). Our factors are not susceptible to either critique. First, as seen in Figure 1, all of our factors (excluding the MKT factor) show positive performance since 1987 and for most subperiods as well. There is a clear, consistent, upward trend. Secondly, our factors are tradeable, such that if, in the future they perform poorly, an astute investor can short them to obtain a positive return.

II. Asset pricing test criteria, portfolios, and alternative models

Although the literature provides compelling evidence for the factors (in some form) in our factor model, it has not been tested in the form we propose. To fully test our factor model, we now describe our test criteria, the test asset portfolios employed, and the alternative models against which we evaluate the performance of our model.

A factor model that prices commodity returns should have an intercept of zero, on average: i.e., there should be zero alpha, both economically and statistically. For each portfolio of commodity futures, we report the alpha and t-statistics. For the entire sorted set of portfolios, we also report the GRS p-value from Gibbons, Ross, and Shanken (1989), which tests the joint hypothesis that all estimates of alpha are zero for the set of test portfolios. We additionally require that a factor model pricing commodity returns should have a high R^2: i.e. it explains a large amount of the variation in the test asset portfolios (Bollen (2013)). We report adjusted R^2

for each test portfolio since the benchmark models have different numbers of explanatory factors.

Our test assets are portfolios sorted on basis and on momentum. Because the literature has converged on basis and momentum as the two key characteristics explaining commodity returns, we focus on those two for brevity.[12] Table IV provides summary statistics for the test portfolios. Panel A shows spot premia sorts on basis and momentum. Both portfolio sets are monotonically ordered with statistically significant high-minus-low portfolio returns of 0.84%, with t-statistics of 2.7 (Basis) and 2.56 (Momentum). Panel B shows term premia sorts on basis for two-, four-, and six-month term premia. They do not show the same ordering, but almost all portfolios show positive alpha significantly different from zero.

For comparison, we run a horse race against two candidate models in the literature. The the *FH model* is one of the most popular models in the hedge fund/CTA literature. Since we are focused on commodities and not the universe of possible assets available to hedge funds, our *FH model* only includes their primitive trend following factor for commodities, called FHCOM in Table III. Other factors related to interest rates, emerging market equities, equity options, currencies, etc. do not apply and only add noise. We obtain this factor from David Hsieh's website.

The second model we have called the *BGR model*, which we now define more rigorously. This is a three factor model containing the MKT, HML, and cross-sectional MOM factor with two groups, split at the median. This model is asserted (without testing) in Bhardwaj, Gorton, and Rouwenhorst (2014) as an appropriate CTA benchmark, along with additional factors for equity options and currencies. Again, we omit these unrelated, latter factors since our focus is on

[12] Szymanowska et al. (2014) rigorously test a variety of other test assets based on other fundamentals, such as inflation, liquidity, and open interest. These additional tests do not materially change their conclusions.

commodities. Bakshi, Gao Bakshi, and Rossi (2014) uses this same factor but uses the top and bottom 5 set of commodities to define H and L, respectively, thus setting N_g equal to 5. Since both sets of authors result can be abbreviated as “BGR”, we use the abbreviation collectively to refer to both of these models, using the MOM definition from Bhardwaj, Gorton, and Rouwenhorst (2014) with N_g equal to 10.

Finally, to recap our model in clear contrast to these two alternatives, our four factors are as follows: a MKT factor, a TSMOM factor, and two basis term premia factors, H_{term} and L_{term}. Thus, we differ from the FH model by including MKT, H_{term} and L_{term}, and using a TSMOM factor to capture trend-following behavior instead of the lookback option approach in Fung and Hsieh (2001). We differ from the BGR model by using TSMOM instead of MOM, and omitting the single HML in favor of H_{term} and L_{term}, and are similar in that we both use the MKT factor.

III. Pricing commodity risk premia

We now present results for our asset pricing tests, starting with spot risk premia. We follow that with a brief discussion on the redundancy of the spot basis factor (HML) and cross-sectional momentum factor (MOM) versus the market factor (MKT) and time-series momentum factor (TSMOM). The final section discusses asset pricing tests on term premia and again revisits the question of parsimony utilizing the full four-factor model.

A. Pricing tests on spot premia portfolios

Table V presents results for each of the three candidate models tested: the Four-factor model, the BGR model, and the FH model. Tests are on basis-sorted portfolios B1-B4 in Panel A and momentum-sorted portfolios M1-M4 in Panel B. Our four-factor model performs well, pricing all test assets in both Panel A and Panel B. Alphas range from 5 bps to 19 bps in absolute value and t-statistics range from 0.34 to 1.28 in absolute value. R^2 range from 0.61 to 0.79, and the GRS p-values are 67.20% and 71.01% for Panels A and B, respectively.

Next, we turn to the BGR model. This model successfully prices all of the test asset portfolios. In Panel A, monthly alpha ranges from 1 bps to 3 bps in absolute value, and t-statistics range from 0.01 to 0.22 in absolute value. Adjusted R^2 is between 0.71 and 0.75. The GRS p-value is 99.75%, strongly supporting the hypothesis that joint alpha is statistically zero. In Panel B, the results are not as strong in magnitude, but statistically give the same conclusion. Monthly alphas range from 3 bps to 19 bps in absolute value, and t-statistics from 0.32 to 1.57 in absolute value, thus indicating that alphas are statistically zero. The adjusted R^2 varies from 0.59 up to 0.90, and the GRS p-value is 54.42%, again supporting the hypothesis that joint alpha is zero.

The FH model can price most of the individual portfolios (6 of 8), but finds significant and economically large negative monthly alpha in portfolios B1 (-68 bps monthly) and M1 (-61 bps monthly). It also rejects the null hypothesis of each portfolio's alpha jointly set to zero with the GRS test, since the p-value is less than 5%. Most importantly, the adjusted R^2 for all portfolios is 0 to two significant figures. This means that this factor model has zero explanatory power of these portfolio returns.

Overall, our Four-factor model performs very well. The magnitudes of alpha are somewhat larger than BGR, as are the t-statistics of alpha. The adjusted R^2 and the GRS p-values give somewhat mixed results in terms of ordering. However, it is clear that both the BGR model and the Four-factor model successfully price commodity spot premia very well. The FH model shows clear inadequacies. In unreported results, we show that adding the HML factor to our model can improve its performance in these tests. This improvement is trivial, however, since all tests already show the portfolios to be fully priced. In the next sections, we will give evidence that the HML factor is redundant.

B. Tests for parsimony among spot factors

To derive some insight into these factors and ensure that our model is parsimonious, we next discuss asset pricing tests on the individual factors as well as various pairs of factors. Results are for individual factors are displayed in Table VI for both Basis portfolios (Panel A) and Momentum portfolios (Panel B). Arguably the best performing factor is the HML factor, which by itself sets alpha equal to zero for all test portfolios, with the highest t-statistic (in absolute value) of -1.93 on portfolio M1 in Panel B. The R^2 are rather low for HML, however. We see a similar pattern for both TSMOM and MOM, though each cannot price one portfolio. TSMOM cannot price portfolio B2 in Panel A (t-statistic of -2.01) and MOM cannot price portfolio M1 in Panel B (t-statistic of -2.98). The MKT factor cannot price the 1 and 4 portfolios for both Basis and Momentum, but has the highest R^2 of all factors to a significant degree.

Next we turn to Table VII, which displays results for pairs of factors. Panel A has results of tests on Basis portfolios, Panel B has results for tests on Momentum portfolios. The first thing that stands out is that including the MKT factor is the only way to improve the R^2. Since we saw this in the individual factors, it is not surprising. The pairs with the MKT factor have a *minimum* R^2 of 0.48 (Panel B portfolio M1 MKT and MOM) whereas the pairs without the MKT factor have a *maximum* R^2 of .35 (Panel B portfolios M4 HML and TSMOM or TSMOM and MOM). Thus, to maximize the R^2, we must include the MKT factor. MOM, the cross-sectional momentum factor, does not perform well pairwise. In both Panels, the MKT and MOM pair cannot price the 1 and 4 portfolios. This pair barely passes the GRS test in Panel B, and does not in Panel A. Both HML and TSMOM fare better. In both pairwise tests with the MKT factor, all test portfolios have alphas no different from zero. As expected, HML performs better pricing the basis portfolios, TSMOM fares better pricing the momentum portfolios.

To test for redundancy, we regress each set of factors on the remaining factors in **Error! Reference source not found.**. The first row shows that MOM is redundant compared to the other three, given a t-statistic of 0.33. Row 2 confirms this finding with only MKT and TSMOM as explanatory varibles, giving an intercept of -0.10% and a t-statistic of -0.56. However, row 3 shows that TSMOM is not redundant since it has a positive and significant intercept with t-statistic of 3.07. Removing HML does not change this conclusion regarding TSMOM, as shown in row 4. Finally, row 5 shows that HML is almost redundant with a t-statistic of 1.86, but this is close to the 5% critical value of 1.96, so we reserve judgement for now until we include our two term structure factors based on basis sorts (and thus HML). Recall from Table III that the correlation between HML and H_{term} is -0.35 and between HML and L_{term} is 0.40 (in Panel A). This univariate result previews our later finding that HML is unnecessary.

C. Pricing tests on term premia portfolios

We next turn to pricing term premia in Table IX. This table is a horse race of our Four-factor model vs the BGR model vs the FH model in pricing basis portfolios term premia. Portfolios are composed in exactly the same way as in Table V, but portfolio performance is based on term premia rather than spot premia. This is where our Four-factor model shines, performing significantly better than the other two. First, the magnitudes of alpha are very low, ranging from 0 bps (Panel C, portfolio B2) to -9 bps (Panel C, portfolio B3). It prices 10 of 12 portfolios across Panels A, B, and C. The GRS test shows in both Panel B and Panel C that all four portfolios jointly have alpha of zero. Panel A, the two-month term premia, does not fare well on the GRS test due to portfolio B4. But even here, the measured alpha is just 8 bps monthly.

In contrast, the BGR model and FH model both perform poorly pricing term premia. The BGR model can price 2 of 12 portfolios (Panel B, B3 and Panel C, B3) and the FH model can price 3 of 12 (portfolio B3 in all three panels). The magnitudes of alpha are much higher, as high as 36 bps monthly for the BGR model (Panel C, B4) and 45 bps monthly for the FH model (Panel C, B1). The GRS test gives a p-value of zero to two significant figures for all groups of portfolios. The BGR model gives R^2 substantially lower than our Four-factor model and the FH model has R^2 that are zero for all portfolios.

Overall, these results provide robust evidence that our Four-factor model performs the best of these three models given our test assets and test criteria. Next, we do some additional checks to be sure our model is as parsimonious as possible. In particular, would a single HML_{term} factor, computed as $H_{term} - L_{term}$, suffice to price term premia? We answer this question in Table X by comparing the two and how they price the term premia test assets. As can be seen, including a HML_{term} factor as a single fourth factor does not adequately price term premia. The second column replicates the results from Table IX for the Four-factor model including both term factors separately for reference. The combined HML_{term} factor performs about the same as the BGR model, pricing only 2 of 12 portfolios and obtaining a p-value of zero for all GRS tests. The adjusted R^2 are substantially lower. Thus, we reject the single HML_{term} factor in favor of the proposed separate H_{term} and L_{term} factors. This result aligns with a similar test in Szymanowska et al. (2014).

As a final test, we again run regressions of factors on each other to determine if the information collectively contained in subsets fully explains other factors. The results are inTable

XI. In Panel A, we test if the two addition term factors drive out any existing factors. Row 1 and 2 of Panel A shows that the MOM factor is still redundant, confirming earlier results. TSMOM clearly still belongs, as shown by the positive and significant intercept in Row 3. Row 4 now clearly shows the redundancy of HML once H_{term} and L_{term} are included. The intercept of 12 bps and t-statistic of 0.76 clearly show that HML is not necessary to explain alpha. Finally, in Panel B, we see that both H_{term} and L_{term} are not redundant. The intercept is positive and significant for a wide variety of explanatory variables.

IV. Conclusion

We establish and test a monthly, implemental, four-factor model of commodity returns. Our factors include a market factor, a time series momentum factor, and H_{term} and L_{term} factors sorted on basis and containing returns to calendar spreads to price the term premia inherent in commodity futures. This model outperforms the existing models of commodity returns existing in the literature Bhardwaj, Gorton, and Rouwenhorst (2014), Bakshi, Gao Bakshi, and Rossi (2014), Fung and Hsieh (2001) when tested against test portfolios of commodity futures spot and term premia. We believe it to be the most parsimonious model possible.

Considering the recent poor performance of passive investments in commodities, our benchmark for active commodity management is a timely addition to the literature. In spite of recent performance, large investors like CalSTRS (California State Teachers Retirement System) are still willing to invest in commodities.[13] This benchmark can be used to draw comparisons between passive management, active management via Commodity Trading Advisors, and

---

[13] *As other funds bail on commodities, CalSTRS pursues test drive*, Reuters, Sept 4, 2015. http://www.reuters.com/article/us-california-investment-commodities-idUSKCN0R416J20150904

commodity-focused Exchange Traded Funds as investors large or small consider commodities exposure.

An additional application of our factor model could also be its conversion into an Exchange Traded Fund itself. Since our factors are tradeable, they could easily be incorporated into a “smart beta” style commodity ETF that uses algorithms to rebalance the portfolio based on market data and factor construction. Such a product would likely be well received in the current marketplace given the prevalence of smart beta products and the desire for commodity exposure.

References

Asness, Clifford S, Tobias J Moskowitz, and Lasse Heje Pedersen, 2013, Value and Momentum Everywhere, *Journal of Finance* 68, 929–985.

Bakshi, Gurdip, Xiaohui Gao Bakshi, and Alberto G Rossi, 2014, Understanding the Sources of Risk Underlying the Cross-Section of Commodity Returns, *SSRN Working Paper*.

Baltas, Akindynos-Nikolaos, and Robert Kosowski, 2012, Momentum Strategies in Futures Markets and Trend-following Funds, *SSRN Working Paper*, 1–60.

Bhardwaj, Geetesh, Gary B Gorton, and K Geert Rouwenhorst, 2014, Fooling Some of the People All of the Time: The Inefficient Performance and Persistence of Commodity Trading Advisors, *Review of Financial Studies* 27, 3099–3132.

Bhardwaj, Geetesh, Gary B Gorton, and K Geert Rouwenhorst, 2015, Facts and Fantasies About Commodity Futures Ten Years Later, *SSRN Working Paper*.

Bollen, Nicolas P B, 2013, Zero-R2 Hedge Funds and Market Neutrality, *Journal of Financial and Quantitative Analysis* 48, 519–547.

Bollen, Nicolas P B, and Robert E Whaley, 2009, Hedge Fund Risk Dynamics: Implications for Performance Appraisal, *Journal of Finance* 64, 985–1035.

Brennan, Michael J, 1958, The Supply of Storage, *The American Economic Review* 48, 50–72.

Carter, Colin A, Gordon C Rausser, and Andrew Schmitz, 1983, Efficient Asset Portfolios and the Theory of Normal Backwardation, *Journal of Political Economy* 91, 319–331.

Cheng, Ing-Haw, and Wei Xiong, 2014, Financialization of Commodity Markets, *Annual Review of Financial Economics* 6, 419–441.

Cooper, Rick, 1993, Risk premia in the futures and forward markets, *Journal of Futures Markets* 13, 357–371.

Dusak, Katherine, 1973, Futures Trading and Investor Returns: An Investigation of Commodity Market Risk Premiums, *Journal of Political Economy* 81, 1387–1406.

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

Fama, Eugene F, and Kenneth R French, 1987, Commodity Futures Prices: Some Evidence on Forecast Power, Premiums, and the Theory of Storage, *The Journal of Business* 60, 55–73.

Fama, Eugene F, and Kenneth R French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F, and Kenneth R French, 2015, A five-factor asset pricing model, *Journal of*

Financial Economics 116, 1–22.

Fuertes, Ana-Maria, Joëlle Miffre, and Georgios Rallis, 2010, Tactical allocation in commodity futures markets: Combining momentum and term structure signals, *Journal of Banking and Finance* 34, 2530–2548.

Fung, William, and David A Hsieh, 1997, Survivorship Bias and Investment Style in the Returns of CTAs, *The Journal of Portfolio Management* 24, 30–41.

Fung, William, and David A Hsieh, 2000, Performance Characteristics of Hedge Funds and Commodity Funds: Natural vs. Spurious Biases, *Journal of Financial and Quantitative Analysis* 35, 291–307.

Fung, William, and David A Hsieh, 2001, The risk in hedge fund strategies: theory and evidence from trend followers, *Review of Financial Studies* 14, 313–341.

Gibbons, Michael R, Stephen A Ross, and Jay Shanken, 1989, A Test of the Efficiency of a Given Portfolio, *Econometrica* 57, 1121–1152.

Gorton, Gary B, and Geert Rouwenhorst, 2006, Facts and Fantasies about Commodity Futures, *Financial Analysts Journal* 62, 47–68.

Gorton, Gary B, Fumio Hayashi, and K Geert Rouwenhorst, 2012, The Fundamentals of Commodity Futures Returns, *Review of Finance* 17, 35–105.

Jegadeesh, Narasimhan, and Sheridan Titman, 2011, Momentum, *Annual Review of Financial Economics* 3, 493–509.

Kaldor, Nicholas, 1939, Speculation and Economic Stability, *The Review of Economic Studies* 7, 1.

Kosowski, Robert, Narayan Y Naik, and Melvyn Teo, 2007, Do hedge funds deliver alpha? A Bayesian and bootstrap analysis, *Journal of Financial Economics* 84, 229–264.

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

Moskowitz, Tobias J, Yao Hua Ooi, and Lasse Heje Pedersen, 2012, Time series momentum, *Journal of Financial Economics* 104, 228–250.

Newey, Whitney K, and Kenneth D West, 1987, A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703–708.

Roll, Richard, 1978, Ambiguity when Performance is Measured by the Securities Market Line, *Journal of Finance* 33, 1051–1069.

Routledge, Bryan R, Duane J Seppi, and Chester S Spatt, 2000, Equilibrium Forward Curves for

Commodities, *Journal of Finance* 55, 1297–1338.

Rouwenhorst, K Geert, and Ke Tang, 2012, Commodity Investing, *Annual Review of Financial Economics* 4, 447–467.

Szymanowska, Marta, Frans A De Roon, Theo E Nijman, and Rob ven den goorbergh, 2014, An Anatomy of Commodity Futures Risk Premia, *Journal of Finance* 69, 453–482.

Working, Holbrook, 1949, The Theory of Price of Storage, *The American Economic Review* 39, 1254–1262.

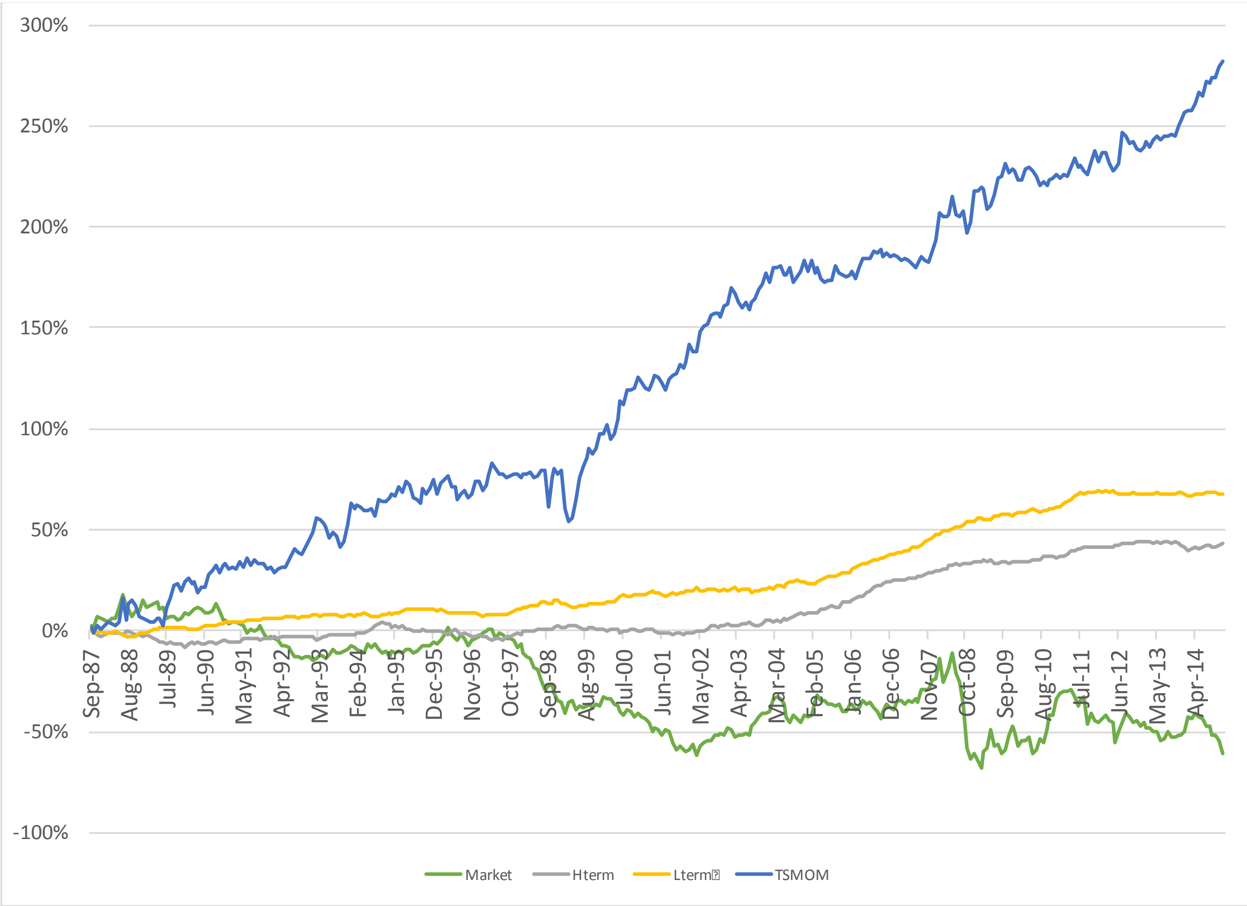

Figure 1: Cumulative returns of four factors
Plotted is the cumulative return of the four factors. Market is an equally-weighted portfolio of all commodity futures sport returns. H_{term} and L_{term} are portfolios sorted on basis, with high defined as above median and low defined as below median set of high and low commodities, and reflect the equal-weighted average return to calendar spreads that are long 2-, 4-, and 6-month maturity futures and short the spot contract. Cumulative returns are computed by summing returns starting in September 1987, and thus remove the effect of compounding.

Table I
List of commodity trading advisor categories

Listed are the categories for CTAs available in Barclay Hedge, along with counts of unique funds included in each category. They are summarized by strategy. Data from Barclay Hedge obtained monthly from December 2006 through December 2014 and verified to be free of the graveyard bias identified in Bhardwaj, Gorton, and Rouwenhorst (2014). CTA summary includes both active and dead funds.

Categories	**Unique Funds in Dataset**	**% of Funds in Dataset**
Commodities		
Agricultural	87	2.9%
Energy	49	1.7%
Financial/Metals	435	14.7%
Commodities Subtotal:	***571***	***19.3%***
Currency	421	14.2%
Interest Rates	29	1.0%
Stock Index	184	6.2%
Other Futures Subtotal:	***634***	***21.4%***
Arbitrage	50	1.7%
Discretionary	53	1.8%
Option Strategies	153	5.2%
Systematic	76	2.6%
Diversified	1,196	40.4%
No Category	228	7.7%
General Strategies Subtotal:	***1,756***	***59.3%***
Total Funds in Dataset:	2,961	100%

Table II
List of commodities included in study

Column 1 is the name, column 2 is the exchange on which they are traded, column 3 is the Bloomberg (BB) symbol, and column 4 is the Commodity Systems, Inc. (CSI) symbol.

Name	Exchange	BB symbol	CSI data symbol
Corn	CBOT	C	C2
Rice Rough	CBOT	RR	RR2
Lumber	CME	LB	LB
Wheat	CBOT	W	W2
Oats	CBOT	O	O2
Coffee	ICE-US	KC	KC2
Cocoa	ICE-US	CC	CC2
Cotton	ICE-US	CT	CT2
Hogs Lean	CME	LH	LH
Soybean Oil	CBOT	BO	BO2
Orange Juice	ICE-US	OJ	OJ2
Silver	COMEX	SI	SI2
Gold	COMEX	GC	GC2
Soybeans	CBOT	S	S2
Feeder Cattle	CME	FC	FC
Cattle Live	CME	LC	LC
NY Harbor ULSD	NYMEX	HO	HO2
Crude Oil Light	NYMEX	CL	CL2
Soybean meal	CBOT	SM	SM2
Copper HG	COMEX	HG	HG2
Gasoline RBOB	NYMEX	XB	RB2

Table III
Performance measurement model summary statistics

This table reports summary statistics and the cross-correlations of five candidate factors to explain the cross section of commodity risk premia. The market factor is an equally weighted average of all futures contracts. The high-minus-low (HML) factor is the difference between the above- and below-median portfolios sorted on spot basis. The time series momentum (TSMOM) is an equally weighted return of commodities with positive 12-month trailing return less those with a negative trailing 12-month return. H_{term} and L_{term} are constructed from three equally weighted calendar spread portfolios of two, four, and six months, split on the median basis for high and low. Cross-sectional momentum (MOM) is an equally weighted return of commodities with above median 12-month trailing return less those with below median 12-month trailing return. FHCOM is the Fung-Hsieh primitive trend following commodity factor from David Hsieh's website. Panel B is subset to start in 1994 because that is when FHCOM is first available.

Panel A: September 1987 - December 2014

				Cross-Correlations					
Factor	**Monthly Excess Return**	**Std Dev**	**T-stat for mean = 0**	**MKT**	**HML**	**TSMOM**	**H_{term}**	**L_{term}**	**MOM**
MKT	-0.19%	3.46%	-0.97	1.00					
HML	0.60%	3.46%	3.12	0.03	1.00				
TSMOM	0.86%	4.39%	3.55	0.20	0.39	1.00			
H_{term}	0.13%	0.73%	3.22	-0.19	-0.35	-0.19	1.00		
L_{term}	0.21%	0.63%	5.94	-0.36	0.40	0.08	0.11	1.00	
MOM	0.84%	5.95%	2.56	0.13	0.51	0.80	-0.23	0.20	1.00

Panel B: January 1994 - December 2014

				Cross-Correlations						
Factor	**Monthly Excess Return**	**Std Dev**	**T-stat for mean = 0**	**MKT**	**HML**	**TSMOM**	**H_{term}**	**L_{term}**	**MOM**	**FHCOM**
MKT	-0.21%	3.72%	-0.90	1.00						
HML	0.64%	3.33%	3.05	0.05	1.00					
TSMOM	0.87%	4.54%	3.03	0.21	0.43	1.00				
H_{term}	0.18%	0.73%	3.86	-0.16	-0.34	-0.15	1.00			
L_{term}	0.24%	0.67%	5.65	-0.38	0.38	0.09	0.17	1.00		
MOM	0.85%	5.97%	2.26	0.12	0.51	0.79	-0.14	0.20	1.00	
FHCOM	-0.24%	14.21%	-0.26	0.01	0.02	0.04	-0.08	0.01	0.05	1.00

Table IV

Portfolio performance, sorted on basis and momentum: September 1987 to December 2014

This table reports summary statistics for portfolios sorted based on basis and momentum. Basis is computed as the log of the ratio of the nearest-dated contract and next-nearest dated contract. Momentum is computed based on excess return over the past year. Each portfolio contains on average of five commodities and is rebalanced monthly. Monthly excess return is computed as the return on the futures contract, without adding in any returns because of collateral reinvestment, and as such is net of the risk-free rate. Panel A shows spot premia, Panel B term premia.

Panel A

Basis Portfolios	Monthly Excess Return	Standard Deviation	t-stat for Mean = 0	Momentum Portfolios	Monthly Excess Return	Standard Deviation	t-stat for Mean = 0
B1 (bottom)	-0.58%	4.54%	-2.30	M1 (bottom)	-0.65%	4.70%	-2.50
B2	-0.42%	4.53%	-1.68	M2	-0.38%	4.09%	-1.67
B3	-0.03%	4.33%	-0.15	M3	0.04%	4.09%	0.20
B4 (top)	0.26%	4.84%	0.97	M4 (top)	0.19%	5.22%	0.67
B4-B1	0.84%	5.61%	2.70	M4-M1	0.84%	5.95%	2.56

Panel B

Basis Portfolios	Monthly Excess Return	Standard Deviation	t-stat for Mean = 0
Panel A. Two Month Term Premia			
B1 (bottom)	0.18%	0.63%	5.17
B2	0.10%	0.48%	3.86
B3	0.07%	0.50%	2.38
B4 (top)	0.21%	0.98%	3.78
Panel B. Four Month Term Premia			
B1 (bottom)	0.26%	1.02%	4.59
B2	0.16%	0.74%	4.00
B3	0.05%	0.74%	1.15
B4 (top)	0.23%	1.51%	2.76
Panel C. Six Month Term Premia			
B1 (bottom)	0.37%	1.37%	4.95
B2	0.17%	0.95%	3.23
B3	0.01%	0.91%	0.28
B4 (top)	0.28%	1.85%	2.70

Table V

Asset pricing tests of spot returns: comparing factor models

Data include 21 commodities' monthly spot returns from September 1987 to December 2014. This table reports asset pricing tests for spot returns when futures are sorted on basis and momentum. Basis is computed as the log of the ratio of the nearest-dated contract and next-nearest dated contract. Momentum is computed based on excess return over the past year. Each portfolio contains, on average, five commodities and is rebalanced monthly. Monthly excess return is net of the risk-free rate (no collateral reinvestment). The *Four-factor model* includes two spot factors (MKT and TSMOM) and two basis term premia factors (H_{term} and L_{term}) to form our our four-factor model. The *BGR model* contains the three commodity factors (MKT, HML and MOM) in Bhardwaj, Gorton, and Rouwenhorst (2014), similar to Bakshi, Gao Bakshi, and Rossi (2014). The *FH model* includes the commodity primitive trend-following factor from Fung and Hsieh (2001). The *FH model* is tested on data starting in 1994 because that is when those factors are first available. Spot returns are based on the nearest-dated contract. T-statistics are computed based on standard errors with a Newey-West correction of 12 lags. GRS p-values are computed as in Gibbons, Ross, and Shanken (1989).

Four-factor Model				***BGR Model***				***FH Model***			
	Alpha	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**
Panel A. Basis portfolios											
B1	0.07%	0.38	0.61	B1	0.00%	0.01	0.74	B1	-0.68%	-2.19	0.00
B2	-0.20%	-1.59	0.68	B2	0.00%	-0.01	0.74	B2	-0.41%	-1.51	0.00
B3	-0.05%	-0.34	0.69	B3	-0.03%	-0.22	0.75	B3	0.07%	0.23	0.00
B4	0.19%	1.28	0.65	B4	0.03%	0.22	0.71	B4	0.13%	0.43	0.00
GRS(p-value)			67.20%	GRS(p-value)			99.97%	GRS (p-value)			2.98%
Panel B. Momentum portfolios											
M1	0.13%	1.03	0.74	M1	-0.03%	-0.32	0.88	M1	-0.61%	-2.39	0.00
M2	-0.15%	-0.98	0.61	M2	-0.17%	-1.28	0.59	M2	-0.48%	-1.67	0.00
M3	0.10%	0.83	0.72	M3	0.19%	1.57	0.71	M3	-0.04%	-0.11	0.00
M4	-0.11%	-0.75	0.79	M4	-0.03%	-0.32	0.90	M4	0.25%	0.82	0.00
GRS(p-value)			71.01%	GRS(p-value)			54.42%	GRS (p-value)			4.71%

Table VI

Asset pricing tests of spot returns: individual factors

Data include 21 commodities' monthly spot returns from September 1987 to December 2014. This table reports asset pricing tests on individual factors for spot returns when futures are sorted on basis (Panel A) and momentum (Panel B). Basis is computed as the log of the ratio of the nearest-dated contract and next-nearest dated contract. Momentum is computed based on excess return over the past year. Each portfolio contains, on average, five commodities and is rebalanced monthly. Monthly excess return is net of the risk-free rate (no collateral reinvestment). The market factor is an equally weighted average of all futures contracts. The high-minus-low (HML) factor is the difference between the above- and below-median portfolios sorted on spot basis. Time series momentum (TSMOM) is an equally weighted return of commodities with positive 12-month trailing return less those with negative trailing 12-month return. Cross-sectional momentum (MOM) is an equally weighted return of commodities with above median 12-month trailing return less those with below median 12-month trailing return. Spot returns are based on the nearest-dated contract. T-statistics are computed based on standard errors with a Newey-West correction of 12 lags.

Panel A: Basis Portfolios

	Alpha	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**
High-Minus-Low (HML)				*Market (MKT)*			
B1	-0.20%	-0.78	0.23	B1	-0.41%	-2.08	0.49
B2	-0.20%	-0.95	0.07	B2	-0.22%	-1.82	0.65
B3	-0.25%	-0.96	0.08	B3	0.16%	1.24	0.68
B4	-0.14%	-0.65	0.23	B4	0.44%	2.32	0.50
GRS(p-value)			87.58%	GRS(p-value)			4.37%
Time series momentum (TSMOM)				*Cross-sectional momentum (MOM)*			
B1	-0.51%	-1.66	0.00	B1	-0.57%	-1.88	0.00
B2	-0.48%	-2.01	0.00	B2	-0.43%	-1.88	0.00
B3	-0.22%	-0.79	0.04	B3	-0.06%	-0.23	0.01
B4	-0.10%	-0.46	0.14	B4	0.23%	1.02	0.01
GRS(p-value)			21.37%	GRS(p-value)			3.63%

Panel B: Momentum Portfolios

	Alpha	T stat	Adj-R2		Alpha	T stat	Adj-R2
High-Minus-Low (HML)				*Market (MKT)*			
M1	-0.40%	-1.93	0.09	M1	-0.48%	-3.96	0.47
M2	-0.36%	-1.49	0.00	M2	-0.21%	-1.53	0.59
M3	0.01%	0.03	0.00	M3	0.23%	1.96	0.71
M4	-0.09%	-0.35	0.09	M4	0.41%	2.24	0.59
GRS(p-value)			22.05%	GRS(p-value)			3.27%
Time series momentum (TSMOM)				*Cross-sectional momentum (MOM)*			
M1	-0.32%	-1.42	0.12	M1	-0.64%	-2.98	0.00
M2	-0.39%	-1.44	0.00	M2	-0.40%	-1.56	0.01
M3	-0.20%	-0.76	0.09	M3	0.03%	0.12	0.00
M4	-0.41%	-1.78	0.34	M4	0.15%	0.68	0.02
GRS(p-value)			40.99%	GRS(p-value)			2.60%

Table VII
Asset pricing tests of spot returns: pairs of factors

Data include 21 commodities' monthly spot returns from September 1987 to December 2014. This table reports asset pricing tests on pairs of factors for spot returns when futures are sorted on basis (Panel A) and momentum (Panel B). Basis is computed as the log of the ratio of the nearest-dated contract and next-nearest dated contract. Momentum is computed based on excess return over the past year. Each portfolio contains, on average, five commodities and is rebalanced monthly. Monthly excess return is net of the risk-free rate (no collateral reinvestment). The market factor is an equally weighted average of all futures contracts. The high-minus-low (HML) factor is the difference between the above- and below-median portfolios sorted on spot basis. Time series momentum (TSMOM) is an equally weighted return of commodities with positive 12-month trailing return less those with negative trailing 12-month return. Cross-sectional momentum (MOM) is an equally weighted return of commodities with above median 12-month trailing return less those with below median 12-month trailing return. Spot returns are based on the nearest-dated contract. T-statistics are computed based on standard errors with a Newey-West correction of 12 lags.

Panel A: Basis Portfolios

	Alpha	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**
MKT and HML				*TSMOM and MOM*			
B1	-0.01%	-0.09	0.74	B1	-0.51%	-1.62	0.00
B2	0.01%	0.09	0.74	B2	-0.49%	-2.00	0.00
B3	-0.05%	-0.38	0.75	B3	-0.23%	-0.82	0.05
B4	0.05%	0.38	0.71	B4	-0.11%	-0.50	0.15
GRS(p-value)			99.73%				21.80%
MKT and TSMOM				*MKT and MOM*			
B1	-0.19%	-1.06	0.54	B1	-0.38%	-1.97	0.49
B2	-0.13%	-1.23	0.66	B2	-0.22%	-1.78	0.65
B3	0.11%	0.83	0.68	B3	0.15%	1.23	0.68
B4	0.19%	1.09	0.56	B4	0.43%	2.20	0.50
GRS(p-value)			56.36%	GRS(p-value)			5.81%
HML and TSMOM				*HML and MOM*			
B1	-0.28%	-1.08	0.24	B1	-0.20%	-0.79	0.23
B2	-0.33%	-1.44	0.11	B2	-0.21%	-1.01	0.08
B3	-0.32%	-1.14	0.09	B3	-0.26%	-0.99	0.08
B4	-0.28%	-1.30	0.27	B4	-0.15%	-0.69	0.23
GRS(p-value)			65.25%	GRS(p-value)			86.11%

Panel B: Momentum Portfolios

	Alpha	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**
MKT and HML				*TSMOM and MOM*			
M1	-0.21%	-1.44	0.57	M1	-0.32%	-1.40	0.12
M2	-0.18%	-1.39	0.60	M2	-0.40%	-1.47	0.00
M3	0.21%	1.66	0.71	M3	-0.20%	-0.75	0.09
M4	0.14%	0.93	0.68	M4	-0.42%	-1.82	0.35
GRS(p-value)			32.50%	GRS(p-value)			38.27%
MKT and TSMOM				*MKT and MOM*			
M1	0.02%	0.21	0.72	M1	-0.45%	-3.78	0.48
M2	-0.08%	-0.62	0.61	M2	-0.22%	-1.56	0.59
M3	0.11%	1.03	0.72	M3	0.23%	1.93	0.70
M4	-0.08%	-0.56	0.78	M4	0.39%	2.13	0.60
GRS(p-value)			88.92%	GRS(p-value)			4.19%
HML and TSMOM				*HML and MOM*			
M1	-0.24%	-1.07	0.15	M1	-0.40%	-1.89	0.09
M2	-0.38%	-1.44	0.00	M2	-0.37%	-1.52	0.01
M3	-0.17%	-0.64	0.09	M3	0.00%	0.00	0.00
M4	-0.46%	-1.95	0.35	M4	-0.10%	-0.43	0.10
GRS(p-value)			32.69%	GRS(p-value)			21.80%

Table VIII

Tests for redundancy among factors in factor model – spot premia

Data include 21 commodities' monthly spot returns from September 1987 to December 2014. This table reports results from regressions of the factors on each other to test for redundancy. If a factor has an intercept no different from zero, then it is redundant. Momentum (MOM) is a cross-sectional momentum factor not used in our final model. It is defined as the top-quartile portfolio less the bottom-quartile portfolio of commodities sorted on the previous 12 months of spot returns. The market (MKT) factor is an equally weighted average of all futures contracts. The high-minus-low (HML) factor is the difference between the above- and below-median portfolios sorted on spot basis. Basis is computed as the log of the ratio of the nearest-dated contract and next-nearest dated contract. Time series momentum (TSMOM) is an equally weighted return of commodities with positive 12-month trailing return less those with negative trailing 12-month return. Spot returns are based on the nearest-dated contract. T-statistics are computed based on standard errors with a Newey-West correction of 12 lags.

Dependent Variable	Intercept	t-statistic	Adj-R2	Independent Variables
MOM	-0.23%	-1.34	0.68	MKT, HML and TSMOM
MOM	-0.10%	-0.56	0.63	MKT and TSMOM
TSMOM	0.40%	3.07	0.64	MKT, HML, MOM
TSMOM	0.61%	3.31	0.19	MKT and HML
HML	0.31%	1.86	0.15	MKT and TSMOM

Table IX

Asset pricing tests of term premia: basis portfolios

Data include 21 commodities' 2-, 4-, and 6-month term premia from September 1987 to December 2014. This table reports asset pricing tests for term premia with holding period returns of two, four, and six months when futures are sorted on basis and computed at sequentially longer dated maturities. Basis is computed as the log of the ratio of the nearest-dated contract and longer dated contract (two, four, or six months later). Each portfolio contains, on average, five commodities and is rebalanced monthly. The *Four-factor model* includes two spot factors (MKT and TSMOM) and two basis term premia factors (H_{term} and L_{term}) to form our our four-factor model. The *BGR model* contains the three commodity factors (MKT, HML and MOM) in Bhardwaj, Gorton, and Rouwenhorst (2014), similar to Bakshi, Gao Bakshi, and Rossi (2014). The *FH model* includes the commodity primitive trend-following factor from Fung and Hsieh (2001). The *FH model* is tested on data starting in 1994 because that is when those factors are first available. T-statistics are computed based on standard errors with a Newey-West correction of 12 lags.

Four-factor model				***BGR model***				***FH model***			
	Alpha	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**
				Panel A. Two Month Term Premia							
B1	0.02%	0.63	0.57	B1	0.13%	2.81	0.14	B1	0.20%	3.31	0.00
B2	0.02%	0.66	0.35	B2	0.09%	2.18	0.03	B2	0.13%	2.82	0.00
B3	0.01%	0.40	0.19	B3	0.07%	2.15	0.01	B3	0.07%	1.83	0.00
B4	0.07%	2.33	0.66	B4	0.24%	3.86	0.06	B4	0.26%	3.38	0.00
GRS (p-value)			0.71%	GRS (p-value)			0.00%	GRS (p-value)			0.00%
				Panel B. Four Month Term Premia							
B1	-0.03%	-1.01	0.68	B1	0.17%	2.39	0.21	B1	0.29%	3.07	0.00
B2	0.01%	0.45	0.43	B2	0.13%	2.73	0.10	B2	0.18%	3.64	0.00
B3	-0.05%	-1.19	0.23	B3	0.06%	1.26	0.03	B3	0.06%	1.05	0.00
B4	0.03%	0.63	0.70	B4	0.30%	3.35	0.10	B4	0.33%	3.02	0.00
GRS (p-value)			28.66%	GRS (p-value)			0.00%	GRS (p-value)			0.00%
				Panel C. Six Month Term Premia							
B1	-0.01%	-0.15	0.67	B1	0.25%	2.91	0.25	B1	0.45%	3.79	0.00
B2	0.00%	-0.12	0.40	B2	0.12%	2.17	0.15	B2	0.17%	2.92	0.01
B3	-0.09%	-2.09	0.23	B3	0.03%	0.58	0.05	B3	0.02%	0.35	0.00
B4	0.05%	0.90	0.71	B4	0.37%	3.51	0.13	B4	0.40%	3.06	0.00
GRS (p-value)			16.49%	GRS (p-value)			0.00%	GRS (p-value)			0.00%

Table X

Tests for HML_{term} factor vs separate H_{term} and L_{term}

Data include 21 commodities' monthly spot returns from September 1987 to December 2014. This table reports asset pricing tests for term premia with holding period returns of two, four, and six months when futures are sorted on basis and computed at sequentially longer dated maturities. It compares a combined HML_{term} factor versus individual H_{term} and L_{term} factors. Basis is computed as the log of the ratio of the nearest-dated contract and longer dated contract (two, four, or six months later). Each portfolio contains, on average, five commodities and is rebalanced monthly. H_{term} is the equally weighted average of the above median portfolio of 2-, 4-, and 6-month calendar spread returns. L_{term} is the same for those below the median. HML_{term} is the difference between the two. T-statistics are computed based on standard errors with a Newey-West correction of 12 lags.

	HML_{term}				***H_{term} and L_{term}***		
	Alpha	**T stat**	**Adj-R2**		**Alpha**	**T stat**	**Adj-R2**
			Panel A. Two Month Term Premia				
B1	0.16%	3.44	0.20	B1	0.02%	0.63	0.57
B2	0.09%	2.48	0.10	B2	0.02%	0.66	0.35
B3	0.08%	2.31	0.07	B3	0.01%	0.40	0.19
B4	0.26%	4.53	0.39	B4	0.07%	2.33	0.66
GRS (p-value)			0.00%	GRS (p-value)			0.71%
			Panel B. Four Month Term Premia				
B1	0.22%	3.11	0.24	B1	-0.03%	-1.01	0.68
B2	0.14%	3.41	0.15	B2	0.01%	0.45	0.43
B3	0.06%	1.41	0.08	B3	-0.05%	-1.19	0.23
B4	0.31%	3.84	0.41	B4	0.03%	0.63	0.70
GRS (p-value)			0.00%	GRS (p-value)			28.66%
			Panel C. Six Month Term Premia				
B1	0.32%	3.57	0.26	B1	-0.01%	-0.15	0.67
B2	0.14%	2.87	0.14	B2	0.00%	-0.12	0.40
B3	0.03%	0.62	0.06	B3	-0.09%	-2.09	0.23
B4	0.38%	3.93	0.43	B4	0.05%	0.90	0.71
GRS (p-value)			0.00%	GRS (p-value)			16.49%

Table XI

Tests for redundancy among factors in four-factor model

This table reports results from regressions of the factors on each other to test for redundancy. If a factor has an intercept no different from zero, then it is redundant. Momentum (MOM) is a cross-sectional momentum factor not used in our final model. It is defined as the top-quartile portfolio less the bottom-quartile portfolio of commodities sorted on the previous 12 months of spot returns. The market (MKT) factor is an equally weighted average of all futures contracts.

The high-minus-low (HML) factor is the difference between the above- and below-median portfolios sorted on spot basis. Basis is computed as the log of the ratio of the nearest-dated contract and next-nearest dated contract. Time series momentum (TSMOM) is an equally weighted return of commodities with positive 12-month trailing return less those with negative trailing 12-month return. H_{term} is the equally weighted average of the above median portfolio of 2-, 4-, and 6-month calendar spread returns. L_{term} is the same for those below the median. Spot returns are based on the nearest-dated contract. T-statistics are computed based on standard errors with a Newey-West correction of 12 lags.

Panel A

Dependent Variable	Intercept	t-statistic	Adj-R2	Independent Variables
MOM	-0.28%	-1.56	0.68	MKT, HML, TSMOM, H_{term}, L_{term}
MOM	-0.23%	-1.31	0.66	MKT, TSMOM, H_{term}, L_{term}
TSMOM	0.63%	3.06	0.18	MKT, HML, H_{term}, L_{term}
HML	0.12%	0.76	0.40	MKT, TSMOM, H_{term}, L_{term}

Panel B

Dependent Variable	Intercept	t-statistic	Adj-R2	Independent Variables
H_{term}	0.17%	3.30	0.15	MKT, HML and TSMOM
L_{term}	0.15%	3.11	0.29	MKT, HML and TSMOM
H_{term}	0.12%	3.08	0.19	MKT, HML and TSMOM, L_{term}
L_{term}	0.12%	3.04	0.33	MKT, HML and TSMOM, H_{term}
H_{term}	0.17%	3.33	0.15	MKT, HML, MOM
L_{term}	0.15%	3.07	0.29	MKT, HML, MOM