Harvesting Commodity Risk Premia

Adrian Fernandez-Perez*, Ana-Maria Fuertes and Joëlle Miffre*****

Abstract:

Recent research in commodity futures pricing has established that long-short strategies based on individual signals (such as the slope of the term structure of commodity futures prices, past performance, or hedging pressure…) offer a better performance than long-only portfolios. However, instead of devoting efforts to comparing the existing *K* individual signals with a view to helping investors to choose ex-ante one signal, our research aims at unifying the literature by combining the strengths of all *K* signals. This paper proposes long-short combined strategies that combine the individual signals. Our results suggest that the combined signals generate more accurate buy/sell recommendations than either one of the *K* individual signals taken in isolation. The combined signals portfolios are also found to better explain the cross-sectional variations in commodity futures returns better than any of the alternatives previously proposed. All in all, we conclude that the combination strategies we propose are found to capture the risk premium of commodity futures better than individual signals.

---

* Research Fellow, Auckland University of Technology, Private Bag 92006, 1142 Auckland, New Zealand. Phone: +64 9 921 9999; Fax: +64 9 921 9940; E-mail: adrian.fernandez@aut.ac.nz

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

*** Professor of Finance, EDHEC Business School, 392 Promenade des Anglais, Nice, France; Tel: +33 (0)4 9318 3255; E-mail: Joelle.Miffre@edhec.edu. Corresponding author.

1. Introduction

Even though the theories of storage and hedging pressure date back to Keynes (1930), Kaldor (1939) and Cootner (1960), the debate surrounding the profitability of long-short strategies in commodity futures markets is still very much alive today. The recent literature documents that the performance of long-short commodity futures portfolios crucially hinges on aspects relating to the slope of the term structure of commodity futures prices, hedging pressure and past performance (Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006; Miffre and Rallis, 2007; Basu and Miffre, 2013). Backwardated commodity futures contracts with high roll-yields, net short hedgers, net long speculators and good past performance are known to outperform contangoed contracts with low roll-yields, net long hedgers, net short speculators and poor past performance.

Aside from these well-known signals, significant spreads in commodity futures returns have also been obtained from long-short strategies based on signals such as value, volatility, open interest, liquidity, dollar beta, inflation beta, or skewness: portfolios of commodities with higher value, higher volatility, higher open interest, lower liquidity, lower dollar betas, higher inflation betas or more negative skewness on average earn more (see e.g., Erb and Harvey, 2006; Gorton *et al*., 2012; Hong and Yogo, 2012; Asness *et al.*, 2013; Szymanowska *et al.*, 2014; Fernandez-Perez *et al.*, 2016). A review of this recent literature is provided by Miffre (2016).

Instead of devoting efforts to comparing the existing K signals with a view to helping investors to choose ex-ante one signal among the many available signals, our research aims at unifying the literature by combining the strengths of all K signals. We propose two approaches that are similar in spirit inasmuch as they both tilt investor's asset allocation towards the commodities that most individual strategies recommend to buy and away from the commodities that most individual strategies recommend to sell. The first approach invests $1/K$ into each of the K individual risk premia. The second approach ranks commodities using a multi-scoring approach. Our results suggest that the combined signals generate more accurate buy/sell recommendations than either one of the K individual signals taken in isolation: they lead to substantial increases in Sharpe, Sortino and Omega ratios relative to a long-only monthly rebalanced portfolio or either one of the individual signal considered in isolation.[1] The $1/K$ and multi-score portfolios are also found to better explain the cross-sectional variations in commodity futures returns better than any of the alternatives previously proposed. All in all, we conclude that the combination strategies we propose are found to capture the risk premium of commodity futures better than individual signals.

We show that the 1/K and multi-score factors allow us to explain the time series and cross sectional characteristics of commodities returns.

2. Futures data and single-score risk premia

A. Futures data

The dataset, provided by Datastream International, comprises of 23 commodity futures spanning all sectors: 12 agriculture products (cocoa, coffee, corn, cotton, frozen concentrated orange juice, oats, rough rice, soybeans, soybean meal, soybean oil, sugar 11 and wheat), 5 energies (gasoline RBOB, heating oil, light sweet crude oil, natural gas and unleaded gas), 2 livestock commodities (feeder cattle, lean hogs), 3 metals (high grade copper, gold, silver 5000) and lumber. Front-end returns are changes in log prices of front-end contract up to one month before maturity; the positions are then rolled to the second-nearest contract. Holding excess returns are defined as the changes in log prices of m^{th}-maturity contract held up to the last day of the month before the front contract matures; the position is then rolled to the m^{th}-maturity contract. Front-end excess returns accrue to investors wishing to capture spot premium, holding excess returns accrue to investors wishing to capture term premium. Spreading excess returns are defined as the difference between front-end excess returns and holding excess returns. The sample spans the period from January 1987 to April 2016. The choices of underlying assets, of m=4 and of time span are dictated by the availability of CFTC data on speculators' and hedgers' positions and by the need to have up to 4 maturities per underlying asset in the forward curve.

This paper focuses on spreading returns measured as the difference in front-end returns and 4^{th} maturity holding returns (m=4). Appendix A provides the results obtained when spreading

---

[1] The Sharpe ratio is measured relative to total risk, the Sortino ratio takes into account downside volatility and the Omega ratio allows for departures from normality and is calculated as the probability weighted ratio of positive versus negative excess returns.

returns are measured with *m*=2 and *m*=3. The definition we use of holding returns follows from Boons and Prado (2015) and assumes that investors hold distant contracts up to the last day of the month before this distant contract is the new front end contract. Assuming instead that investors hold distant contracts up to the last day of the month before these contracts mature (as in Szymanowska *et al*., 2014) yield spreading risk premia that are similar to those reported in this paper. The results are available in the Appendix B. We use the definition of holding excess returns of Boons and Prado (2015) rather than that of Szymanowska *et al*. (2014) as that definition ensures that contracts with different maturities are traded at any point in time as part of the spreads.

B. Single-score risk premia

This section details the economic or theoretical rationale for each of the *K*=11 single-score risk premia, as well as the methodology employed in constructing the long-short single-score portfolios.

The term structure strategy follows from the theory of storage of Kaldor (1939), Working (1949), Brennan (1958) as empirically validated in *e.g.*, Fama and French (1987), Erb and Harvey (2006) or Gorton and Rouwenhorst (2006). The theory of storage relates the slope of the term structure (also called the roll-yield) to inventory levels and to the costs and benefits of owning the physical commodity. The term structure strategy captures the risk premium earned when buying commodities in scarce supply and shorting commodities in abundant supply. The roll-yield signal is defined as $ln(f_{t,Front}) - ln(f_{t,Second})$, where $f_{t,Front}$ is the price of the front-end contract on at the time of portfolio formation *t*. The idea is to buy commodities with higher roll-yields and to short commodities with lower roll-yields.

The hedging pressure strategy follows from the hedging pressure hypothesis of Keynes (1930), Cootner (1960) and Hirshleifer (1988) as empirically validated in *e.g.*, Chang (1985), Bessembinder (1992), De Roon *et al.* (2000) or Basu and Miffre (2013). The hedging pressure hypothesis states that net long (short) speculators demand a risk premium for taking on the risk of price decline (rise) that net short (long) hedgers are willing to get rid of. The hedging pressure strategy attempts to capture this risk premium either by tracking the positions of speculators or by taking positions that are opposite to those of hedgers. The hedging pressure of hedgers signal is defined as $HP_H = \frac{1}{W}\sum_w^W \frac{Short_{H,w}-Long_{H,w}}{Long_{H,w}+Short_{H,w}}$; the hedging pressure of speculators signal is set to $HP_S = \frac{1}{W}\sum_w^W \frac{Long_{S,w}-Short_{S,w}}{Long_{S,w}+Short_{S,w}}$ where $Long_{H,w}$ and $Short_{H,w}$ are the long and short open interests of large commercial traders (hedgers) on observation *w* and likewise for large non-commercial traders (speculators) and W is the total number of observations in the 12-month period preceding portfolio formation.[2] The idea is to buy the commodities with higher HP_H and HP_S and to short commodities with lower HP_H and HP_S.

The momentum strategy follows from Erb and Harvey (2006) and Miffre and Rallis (2007). While less theoretically grounded than the term structure and hedging pressure strategies, the momentum strategy is nonetheless acknowledged to pick up commodities that are prone to perform well according to the theories of storage and hedging pressure (Miffre and Rallis, 2007; Gorton *et al.*, 2012). The momentum signal is set equal to the excess return of a given commodity averaged over the past 12 months. The idea is to buy past winners and to short past losers.

The value strategy follows from the mean reversion literature as discussed in DeBondt and Thaler (1985), Jegadeesh and Titman (2001) or Asness *et al.* (2013). Behavioralists argue that the cognitive errors that investors make when incorporating information into prices lead to first momentum and subsequently mean reversion. As a result, long-term winners are deemed to be future underperformers; and vice versa, long-term losers are deemed to subsequently outperform. We capture mean reversion via a value signal measured at portfolio formation *t* and set equal to $ln\frac{\frac{1}{D}\sum_{d=1}^{D} f_{d,Front}}{f_{t,Front}}$ where *D* is the number of days in a period spanning 4.5 to 5.5 years before portfolio formation. The idea is to act as contrarian by buying long-term losers with higher valuations and shorting long-term winners with lower valuations.

The volatility strategy follows from Dhume (2011), Gorton *et al.* (2012) and Szymanowska *et al.* (2014). Borrowing the setting of the consumption CAPM, Dhume was the first to suggest that high volatility commodities shall outperform because of their higher sensitivity to shocks to durable consumption growth. The logic is that such highly volatile commodities correlate positively with durable consumption growth and thus shall earn more as they fail to hedge

[2] The commitments of traders are reported either at a bi-monthly frequency over the period January 15, 1986 to September 30, 1992 or at a weekly frequency since then.

negative shocks to durable consumption. As in Szymanowska *et al.* (2014), we use as proxy for volatility the coefficient of variation; the latter is measured as $\hat{\sigma}^2/|\hat{\mu}|$ or as the ratio of the variance of front-end excess returns $\hat{\sigma}^2 = \frac{1}{D-1}\sum_{d=1}^{D}\left(r_{d,Front} - \hat{\mu}\right)^2$ to their mean $\hat{\mu} = \frac{1}{D}\sum_{d=1}^{D} r_{d,Front}$ where D is the number of days in the 36 months prior to portfolio formation. The idea is to buy commodities with higher coefficients of variation and to short commodities with lower coefficients of variation.

The open interest strategy follows from Hong and Yogo (2012) and Szymanowska *et al.* (2014). Hong and Yogo (2012) argue that rather than the *net* positions of market participants, it is their *gross* positions that impact commodity prices. The predictive power of open interest over commodity futures prices comes from the fact that the change in open interests is a highly pro-cyclical indicator of future economic activity. The open interest signal is computed as the change in the open interests of each commodity along the whole term structure. The open interest risk premium captures the excess returns that investors demand for holding commodities with higher open interests as opposed to commodities with lower open interests.

The liquidity strategy recognizes that the least liquid an asset is, the higher its expected return. Following Amihud *et al.* (1997), Marshall *et al.* (2012) and Szymanowska *et al.* (2014), the liquidity signal is based on the so-called liquidity ratio (LR, or Amivest measure of liquidity) which is defined as the average daily trading volume associated with a one unit daily change in front-end futures prices; $LR = \frac{1}{D}\sum\frac{\$Volume_d}{|r_d|}$ and D is the number of days in the previous 2-month period. The liquidity risk premium captures the excess returns that investors demand on commodities with lower LR relative to commodities with higher LR.

Erb and Harvey (2006) show that the S&P-GSCI excess returns have a negative and statistically significant sensitivity with the changes in the USD versus major currency index. This indicates that a weaker USD coincides with stronger commodity prices. Following Erb and Harvey (2006) and Szymanowska *et al.* (2014), the FX signal is the slope of 60-month regression of monthly futures returns on the changes in USD versus major currency index. The FX risk premium captures the excess returns that investors demand for holding commodities with lower dollar betas versus those with higher dollar betas.

It has long been acknowledged that commodities provide a hedge against inflation shocks; namely, commodities perform well in inflationary periods (Bodie and Rosansky, 1980; Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006). Following Szymanowska *et al.* (2014), the inflation signal is the slope of 60-month regression of monthly futures returns on inflation shocks.[3] The inflation risk premium captures the excess returns that investors demand for holding commodities with higher inflation betas relative to commodities with lower inflation betas.

Finally the skewness signal follows Fernandez-Perez *et al.* (2016). It is defined as $\frac{\frac{1}{D}\sum_{d=1}^{D}(r_d-\mu)^3}{\sigma^3}$, where D is the number of days in the previous 12-month period. Investors' aversion towards negatively skewed return distribution and preference for positively skewed ones dictate that negatively skewed assets are underpriced and therefore shall outperform. As negatively (positively) skewed commodities present backwardated (contangoed) characteristics, the skewness risk premium also relates to the fundamentals of backwardation and contango as hypothesized by the theories of storage and hedging pressure. The skewness risk premium captures the excess returns that investors demand for holding commodities with highly negative skew as opposed to commodities with highly positive skews.

Let K denote the number of individual signals (i.e., K=11 here) and N the number of commodities available (i.e., N=23 here) at the time of portfolio formation. At the end of each month and for each of the K individual signals, the N commodities are assigned a score or relative rank in the range [1,…,N] based on their expected performance. The commodity that is expected to perform the best is given the highest score of N; the commodity that is expected to offer the second best performance is given the score of N-1; the commodity that is expected to perform the worst is given the score of 1. Each of the K individual strategies buys the quartile with the highest scores (Q4), shorts the quartile with the lowest scores (Q1) and holds the equally weighted long-short positions for one month on a fully-collateralized basis. In practice, this amounts to buying the commodities with either higher roll-yield, higher hedgers' hedging pressure, higher speculators' hedging pressure, higher past excess returns, higher value, higher volatility, higher change in open interest, lower liquidity ratio, lower dollar beta, higher inflation beta or more negative skewness. Simultaneously, the single-score strategies take short positions in the commodities at the other extreme.

---

[3] Inflation is measured as the monthly percentage change in the consumer price index and inflation shocks are set equal to the change in the inflation rate.

C. Summary statistics of the single-score risk premia

The single-score strategies are implemented on front-end contracts (Table 1) or on spreads (Table 2). The signals are the same in both sets of strategies; the only difference comes from the contracts that are being traded: front-end contracts in Table 1 and spreads, defined as the difference in returns between front-end and 4^{th}-maturity contracts, in Table 2. In each table, Panel A presents summary statistics of performance for the Q4-Q1 portfolios and Panel B presents cross risk premia correlations.

Table 1, Panel A confirms that some signals are useful at forecasting the direction of next month front-end futures price changes. In decreasing order of Sharpe ratios, the term structure, skewness, momentum and speculators' hedging pressure strategies, when applied to front-end contracts, perform well both on a risk-adjusted basis and in statistical terms. Even though they had been documented as profitable in the past, other signals perform poorly within the sample and cross-section considered here: these signals include hedgers' hedging pressure, value, volatility, open interest, liquidity, dollar beta and inflation beta. The portfolio with the worst risk-adjusted performance is a long-only monthly-rebalanced portfolio invested in all N commodities (*EW*). This result confirms the well-documented requirement, when modeling commodity futures risk premia, of considering both long and short positions. Table 1, Panel B shows some independent movements across risk premia; for example, the value risk premium (which is contrarian in nature) is found to negatively correlate with most of the other risk premia. The average correlation stands at merely 0.03, highlighting some benefits in terms of diversification that could potentially be attained from combining the signals.

<< Insert Table 1 around here >>

Table 2 reports similar information for spreading strategies. Spreading strategies based on term structure, value and skewness earn positive risk premia that are significant both economically (Sharpe ratio of 0.41 and above) and statistically (at the 5% level). As with the front-end risk premia of Table 1, the spreading risk premium based on the term structure signal stands out since it offers the highest Sharpe ratio (0.57). The other signals prove to produce Q4-Q1 spreading returns that are either marginally significant (Inflation shocks) or not significant at conventional levels. As with the front-end premia in Table 1, Panel B, the average cross risk premia correlation is low at 0.03 and the value spreading strategy is found to negatively correlate with the other spreading risk premia. This suggests that a joint consideration of the signals might successfully diversify the risks associated with each spreading strategy.

<< Insert Table 2 around here >>

3. Combined risk premia: Time series results

A. Methodologies

Does the joint consideration of these *K* individual signals lead to superior performance? The combined strategies here proposed aim at supplying commodity investors with a composite signal that provides relatively more informed and less noisy information about subsequent price changes. This is done by designing combined strategies that tilt investor's asset allocation both towards the commodities that most individual strategies recommend to buy and away from the commodities that most individual strategies recommend to sell. We hope thereby to obtain a modelling of the risk premia that investors demand in commodity markets that is better than the one produced beforehand with the individual risk premia.

To combine the signals we consider two approaches. The most naïve method, hereafter called 1/*K*, consists of equally-weighting the *K* individual strategies and simultaneously netting the buy and sell recommendations that each of the *K* strategies recommends. This approach has been shown to work well in equity markets (see e.g. DeMiguel *et al.*, 2009). Sequential details on the 1/*K* portfolio formation technique are as follows. At the end of each month *t*, we appraise the positions (long, short or neutral) that each single-score strategy recommends for each of the *N* commodities: for each strategy, a quarter of the available commodities enters the long portfolio Q4 and are assigned a score of +1, a quarter enters the short portfolio Q1 and are assigned a score of -1; the remainder are not traded (as they enter the intermediary portfolios Q2 and Q3) and are assigned a score of 0. We then sum up the scores given to each commodity thereby netting the individual positions and ensuring that long (+1) and short (-1) recommendations for a given commodity cancel out in the 1/*K* portfolio. Finally, we hold the net positions on a fully-collateralized basis over the following month ensuring that at month-end 50% of the client's mandate is invested in the longs and 50% is invested in the shorts.[4] In

[4] For instance, suppose 1/*K* mandates to buy {+1} corn, {+1} corn and {+1} cocoa; and sell {-1} crude oil, {-1} coffee and {-1} cocoa, before netting. Given cocoa is bought and sold at the same time, both positions in cocoa are cancelled. So assuming equally weighted, we would invest 50% of the budget in long and 50% in short

effect the so-called 1/*K* strategy heavily invests into the commodities that often appear in the long single-score portfolios Q4 and heavily shorts the commodities that often appear in the short single-score portfolios Q1.

The alternative combining strategy, hereafter called multi-scoring approach, is similar in spirit to the 1/*K* strategy; the only difference is that it is based on a more gradual scoring system. Sequential details on the multi-scoring portfolio formation technique are as follows. At the end of each month *t*, we assign scores to each individual commodity *i* (i=1,…,N) based on each of the individual signal *k* (k=1,…,K). As with the single-scores, the commodity that is expected to perform the best (worst) is given the highest (lowest) score of *N* (1). We then obtain a total score for each commodity as the sum of the *K* individual scores. The multi-score strategy buys the quartile with the highest total score (Q4), shorts the quartile with the lowest total score (Q1) and holds the long-short portfolio on a fully-collateralized basis for one month.

B. Preliminary results

Figure 1 presents the future value of $1 invested at the beginning of the sample period in each of the fully-collateralized long-only and long-short portfolios discussed thus far. The plot reveals the usefulness of combining the signals in terms of both performance and volatility. Figure 2 presents similar information for the spreading strategies. The first and third best performance is obtained with multi-score and 1/*K* strategies, respectively. The value and term structure spreading strategies obtain the best performance in different periods of the sample but they suffer from the drawback of discerning *ex-ante* their remarkable *ex-post* performance. Investors typically do not have such perfect foresight and thus choosing these particular signals might have been difficult, not to say impossible, *ex-ante*.

<< Insert Figures 1 and 2 around here >>

Table 3 presents summary statistics for each quartile (Q1 to Q4), as well as the long-short portfolio (Q4-Q1). Panel A focuses on the front-end risk premia, Panel B on the spreading risk premia. The first set of columns pertains to the 1/*K* portfolios, the second set to the multi-score portfolios. For the sake of comparison, the last column refers to the single-score

---

portfolios which is equal to a weight of 50% in a long position in corn, and 25% in a short position in crude oil and another 25% for another short position in coffee.

strategies that generate the highest Sharpe ratios in Tables 1 and 2; namely, the term structure strategies.

<< Insert Table 3 around here >>

Table 3, Panel A shows that for both the 1/*K* and multi-score quartiles, mean excess returns increase monotonically from portfolio Q1 (made of commodities deemed to underperform) to portfolio Q4 (made of commodities deemed to outperform). The spreads in mean excess returns are positive and significant at the 1% level for both the 1/*K* and multi-score portfolios. Systematically buying the commodities deemed to outperform and selling the commodities deemed to underperform based on the combined signals yield 7.94% (*t*-statistic of 5.29) a year according to the 1/*K* approach and 8.93% (*t*-statistic of 5.03) a year according to the multi-score approach. Compared to the alternative portfolios considered in Table 1, the 1/*K* and multi-score portfolios present the lowest downside volatilities, the lowest 99% values at risk and the highest maximum drawdowns. This suggests that by combining signals investors partially diversify the risk of trading on a wrong signal and thereby reduce the potential losses incurred on the strategy, in addition of limiting the crash risk. As a result, the risk-adjusted performance of the 1/*K* and multi-score portfolios is remarkable: the Sharpe ratios are 0.96 and 0.97; the Sortino ratios stand at 1.75 and 1.77; the Omega ratios equal 2.04 and 2.08 for the 1/*K* and multi-score portfolios, respectively. These ratios exceed by a large margin those reported for the best single-score strategy. As reported in the last column, the Sharpe, Sortino and Omega ratios of the term structure strategy stood at merely 0.65, 1.06 and 1.66, respectively. These results highlight the conclusion that combining signals is a better way to capture the risk premium of commodity futures than treating each signal in isolation. Not only is performance enhanced but also the combination approach eliminates the risk of picking up *ex-ante* a signal that *ex-post* may happen to be unprofitable.

Table 3, Panel B looks into the performance of combined spreading portfolios (Q1 to Q4) and also presents summary statistics on the combined spreading risk premia (Q4-Q1). As in Panel A for the front-end strategies of Table 3, the spreading strategies present mean excess returns that rise monotonically from Q1 to Q4. The yearly Q4-Q1 spreading risk premia equal 2.57% (*t*-statistic of 3.52) for 1/*K* and 2.98% (*t*-statistic of 3.71) for the multi-score. The corresponding Sharpe, Sortino and Omega ratios of the 1/*K* spreading long-short portfolio stand at 0.68, 1.12 and 1.67, respectively. Those obtained with the multi-score approach are of a similar magnitude (0.69, 1.14 and 1.71, respectively).[5] The performance of the combined spreading strategies overcomes to that of the best single-score strategy (based on the term structure signal) for which the Sharpe, Sortino and Omega ratios stand at 0.57, 0.97 and 1.54, respectively. Yet, the single-score strategy can only be noted *ex-post* while the combination approach does not presume perfect insight of investors as to which signal is going to be *ex-post* the best.

Table 4 further investigates the risk and performance of the combined 1/*K* and multi-score risk premia by regressing each of them onto the 12 traditional risk premia (*EW* and single-score) presented in Tables 1 and 2. The first set of columns pertains to the estimated coefficients and *t*-statistics (corrected by Newey-West standard errors) obtained by regressing front-end combined returns on the front-end risk premia of Table 1; the second set of columns present similar information but this time the spreading combined returns are regressed on the spreading risk premia of Table 2. It is interesting to note that all 4 combined risk premia load positively and significantly on the single-score risk premia. This suggests that the combined approaches are useful at capturing the multifaceted risks present in commodity futures markets and thus are capable of measuring the risk premium that investors demand for exposure to these risks. No single long-short risk premium is left out. Further evidence in support of the idea that the combined strategies successfully harvest commodity risk premia is provided by the high adjusted *R*-squares obtained from the regressions (the adjusted *R*-squares range from 72.03% to 97.03%). After of accounting for all the risk premia, the combined strategies do not leave any significant alpha which suggests that the combined approaches are able to successfully pick up the risk premium of commodity futures, leaving no more than what is expected from exposures to the various risk premia.

<< Insert Table 4 around here >>

Investors seek commodity exposure in part to hedge inflation. While long-only commodity portfolios correlate positively with inflation shocks, long-short commodity portfolios tend to be poor inflation hedges; then, short positions eliminate the natural hedge that long positions provide. Part of the asset allocation of the combined strategies consists however of buying inflation hedging commodities while simultaneously shorting commodities that poorly hedge

---

[5] We fail to reject the null hypothesis that the Sharpe ratio of the 1/*K* spreading strategy (at 0.68) equals that of the multi-score spreading strategy (at 0.69) using Opdyke (2007) test for differences in Sharpe ratios.

inflation. It follows that the inflation betas of the combined strategies are expected to be much higher than that of most single-score strategies. Unreported results confirm this intuition: the inflation betas of the combined portfolios (at 1.89 on average) are relatively close to the inflation betas of the *EW* and inflation hedging portfolios (2.83 and 2.51, respectively) and much higher than the inflation betas of the other single-score strategies (0.46 on average). This suggests that our combined approach is quite successful at hedging inflation shocks.

C. Robustness checks

1. Transaction costs

Table 5 studies the net performance of the combined 1/*K* and multi-score strategies after considering round-trip transaction costs of 0.033% and a more conservative of 0.066% (Locke and Venkatesh, 1997). The impact of transaction costs on performance is found to be minimal: performance merely decreases by 24bp (46bp for 0.066%) a year for the front-end strategies and by 42bp (83bp for 0.066%) for the spreading strategies considering transaction costs of 0.033%. We also calculate the level of transaction costs that would set the mean excess returns of each combined strategy equal to zero. The obtained break-even transaction costs equal 1.30% and 0.45% for the combined front-end and spreading strategies respectively and by very far exceed our most conservative estimate of 0.066%. These results highlight the conclusion that the performance obtained thus far will not be eroded by the cost of implementing the trades.

We also look at the turnover of the strategies where turnover is measured as the average sum of the absolute value of the trades across the *N* available commodities. In summary, we observe how the 1/*K* strategy is less trading-intensive than the multi-sort strategy in both front and spreading returns.

<< Insert Table 5 around here >>

2. Data mining

The previous sections have shown that the $1/K$ and multi-sort strategies outperform the individual strategies. However, an important problem when evaluating a large set of trading rules is data mining or snooping. The mining occurs when a set of data is used more than once for inference or model selection. To deal with this problem we implement the test for Superior Predictive Ability (SPA) developed by Hansen (2005).[6] We have m alternative decision rules (point forecasts or trading rules) are compared to a benchmark, where performance is defined in terms of expected loss. The question of interest is whether any alternative trading rule is better than the benchmark. The complexity of this exercise arises from the need to control for the full set of alternatives. The latter leads to a composite null hypothesis and a t-statistic whose (asymptotic) distribution is non-standard and requires bootstrapping.

Our trading rules universe consists in all the possible combinations of signals (1-by-1, 2-by-2, 3-by-3, etc.) using both scoring and $1/K$ with K<11 methodologies, plus the long-only equally-weighted commodity, the $1/K$ and multi-sort portfolios, i.e. $k = 0, 1, 2, \ldots, m$ ($m = 4{,}083$). We will employ three different benchmarks, represented by $k = 0$; 1) the long-short $1/K$, 2) the long-short multi-sort and 3) the long-only EW commodity portfolio. Let $r_{k,t}$ denote the month t returns of trading rule k. The returns are mapped into a ‘‘loss" by means of a linear function, on one hand, and two nonlinear (exponential) functions with different degrees of curvature, on the other hand. The former is $L_{k,t} = max(r_t) - r_{k,t}$, where $max(r_t)$ is the highest return in the full set of strategies $k = 0, 1, \ldots, m$. The two nonlinear functions are $L_{k,t} = 1/exp(\lambda r_{k,t})$ for $\lambda = 1$ and 2. Thus the sample performance of trading rule $k = 1, \ldots, m$ relative to the benchmark is given by $d_{k,t} = L_{0,t} - L_{k,t}$ over $t = 1, \ldots, T$ months. Strategy k is better than the benchmark if and only if $E[d_{k,t}] > 0$ where $E[\cdot]$ denotes expected value. The null hypothesis is that the best of the m active strategies does not outperform the benchmark, i.e. $H0$: $E[d_{k,t}] \leq 0, k = 1, \ldots, m$.[7] What we want to test is if the best of all the possible combinations is statistically better than the $1/K$ and multi-sort strategies which combine all the individual signals. Table 6 reports the results.

---

[6] In the literature, there is another data snooping test, White's (2000) Reality Check (RC), which essentially builds on the same framework than SPA. Compared to RC, the SPA test is based on a studentized test statistic and a sample dependent distribution under the null hypothesis, both of which make it more powerful and less sensitive to the inclusion of poor and irrelevant alternatives. Therefore, we focus only on SPA in this paper.

[7] The implementation is based on the stationary bootstrap of Politis and Romano (1994) based on B = 10,000 pseudo time series $\{d_{k,t}^*\}$ for each k, which are resamples of $d_{k,t}$ constructed by combining blocks with random lengths. The block-length is geometrically distributed according to $q \in [0,1)$ so the expected block-length is $1/q$. Two typical values are used, $q = \{0.2, 0.5\}$, to robustify the results.

<< Insert Table 6 around here >>

With all three loss functions and two q-values, the bootstrap p-values for the null hypothesis that the 'best' trading rule does not outperform either the $1/K$ as the multi-sort strategies unanimously suggest not rejection of the hypotheses at conventional significance levels, thus confirming that the $1/K$ and the multi-sort strategies are successful trading rules. Instead, the bootstrap p-values for the long-only EW portfolio suggest rejection of the null hypotheses at conventional significance levels.

3. Novy-Marx (2016)

Novy-Marx (2016) raises concerns about the usage of strategies that combine multiple signals in the literature. There are two important biases when a researcher employs combined strategies: *selection bias* and *overfitting bias*. At the one extreme, the pure selection bias results when the researcher employs the best performing signal from among multiple candidates, and fails to account for doing so. The pure overfitting bias is the opposite extreme, i.e. when the researcher employs each and every signal considered, even though several of those signals may be purely random. Novy-Marx (2016) also considers a general case where both selection and overfitting biases are present, resulting from combining the best K signals from a set of n candidates, where $1 < K < n$, which is almost as bad as that from using the single best signal out of n^k candidates.

Because of the selection and overfitting biases that result when researchers consider more signals than they employ, and when they sign individual signals to generate positive in-sample returns, the distribution of t-statistics for a multi-signal strategy does not have a standard normal distribution. In other words, it is easy, combining spurious, marginal signals, to generate backtested performance that looks impressive, at least when evaluated using the wrong statistics. Therefore, critical values derived from that distribution consequently cannot be used to draw inferences regarding the significance of the multi-signal strategies. Instead, Novy-Marx (2016) provides two statistics to test the null hypothesis that the mean return of the multi-signal strategies is not significant; one for the 1/*K* case,

$$t_{n,K}^{1/K} = \frac{\sum_{i=1}^{K} t_{n+1-i}}{\sqrt{K}} \tag{1}$$

and another for the multi-score case,

$$t_{n,K}^{Multi-Score} = \sqrt{\sum_{i=1}^{K} t_{n+1-i}^{2}} \tag{2}$$

where t is the vector of t-statistics estimated on the individual signals and sorted on increasing order. These statistics do not follow standard normal distribution, hence, Novy-Marx provides the critical values for two special cases (pure selection bias and pure overfitting bias) and a normal approximation for the general case (both selection and overfitting biases).

Specifically, the critical values for the pure overfitting bias (*K* = *n*)[8] are the following,

$$critical - t_{n,n,p}^{1/K} \approx \left(\sqrt{\frac{2}{\pi}}\right) \cdot \sqrt{n} + \left(\sqrt{1 - \frac{2}{\pi}}\right) \cdot N^{-1}\left(1 - \frac{n \cdot p}{n+1}\right) \tag{3}$$

$$critical - t_{n,n,p}^{Multi-Score} = \sqrt{\Phi_{\chi_n^2}^{-1}(1-p)} \tag{4}$$

where p is the critical value at the {1%, 5%, 10%} level, $N^{-1}(\cdot)$ is the inverse of the cumulative normal distribution, and $\Phi_{\chi_n^2}^{-1}(\cdot)$ denotes the cumulative distribution for the χ^2 with *n* degrees of freedom.

Otherwise, the critical values for the general case (*1* < *K* < *n*) are the following,

[8] Our results do not suffer from pure selection bias so we skip this statistic.

$$critical - t_{n,K,p}^{1/K} \approx \mu_{t_{n,K}^{1/K}} + \sigma_{t_{n,K}^{1/K}} \cdot N^{-1}\left(1 - \frac{K \cdot p}{K+1}\right) \quad (5)$$

$$critical - t_{n,K,p}^{Multi-Score} \approx \sqrt{\mu_{\left(t_{n,K}^{MS}\right)^2} + \sigma_{\left(t_{n,K}^{MS}\right)^2} \cdot N^{-1}\left(1 - \frac{K \cdot p}{K+1}\right)} \quad (6)$$

where

$$\mu_{t_{n,K}^{1/K}} = \sqrt{K} \cdot \lambda_{n,K} \quad (7)$$

$$\sigma^2_{t_{n,K}^{1/K}} = \Sigma_{n,K} - \lambda^2_{n,K} + \frac{K \cdot (n-K) \cdot \left(\lambda_{n,K} - \mu_{n,K}\right)^2}{(K+1) \cdot (n+2)} \quad (8)$$

$$\mu_{\left(t_{n,K}^{MS}\right)^2} = K \cdot \Sigma_{n,K} \quad (9)$$

$$\sigma_{\left(t_{n,K}^{MS}\right)^2} = K \cdot \left(\mu^3_{n,K} \cdot \lambda_{n,K} + 3 \cdot \Sigma_{n,K} - \Sigma^2_{n,K}\right) + \frac{K^2 \cdot (n-K) \cdot \left(\Sigma_{n,K} - \mu^2_{n,K}\right)^2}{(K+1) \cdot (n+2)} \quad (10)$$

and

$$\mu_{n,K} = N^{-1}\left(1 - \frac{K+1}{2 \cdot (n+1)}\right) \quad (11)$$

$$\lambda_{n,K} = 2 \cdot \left(\frac{n+1}{K+1}\right) \cdot n(\mu_{n,K}) \quad (12)$$

$$\Sigma_{n,K} = 1 + \mu_{n,K} \cdot \lambda_{n,K} \quad (13)$$

where $n(\cdot)$ is the standard normal density function.

As usual, the t-statistics in Equations (1) and (2) are significant if these are bigger than the critical values in Equations (3) and (4), respectively, for pure overfitting bias; and Equations (5) and (6), respectively, for a combination of both selection and overfitting biases.

In the Table 7, we report the results of these tests. Assuming pure overfitting bias (*K=n=11*), Panel A of Table 7 reports the Novy-Marx's statistics (Equations 1 and 2) together with the pure overfitting critical values (Equations 3 and 4) at the {1%, 5%, 10%} levels for both front end and spreading returns. Assuming that the number of true candidates signals, *n*, is larger than the *K* signals employed in this paper (*1< K=11 <n*), Panel B of Table 7 reports the

number of candidates signals, *n*, that makes both multi-signal strategies statistically insignificant, i.e., those *n*'s that make Novy-Marx's statistics (Equations 1 and 2) smaller than the general case critical values (Equations 5 and 6) at the {1%, 5%, 10%} levels for both front end and spreading returns.

<< Insert Table 7 around here >>

Panel A of Table 7 shows how both 1/*K* and multi-score are statistically significant after of accounting by pure overfitting bias at the 5% level or better. In the Panel B of Table 7, we see how the multi-score strategy would be statistically significant at the 5% level if the number of true candidate signals used were bigger or equal to than 44 for front end returns and 21 for the spreading returns. Likewise, the 1/*K* strategy would be statistically significant at the 5% level if the number of true candidate signals used were bigger or equal to than 23 for front end returns and 13 for the spreading returns. Knowing that there would not be consensus in the literature about the total number of candidate signals, *n*, that a researcher in commodity futures should employ (see e.g. Daskalaki and Skiadopoulos, 2014; Bakshi et al., 2015; and Miffre, 2016), at the very least, our results show that combining the 11 signals employed in this paper through 1/*K* and multi-scoring procedures would still provide statistically significant mean returns if we would decide (deliberately) to not use about 11 more signals than those employed in this paper (i.e., $K=11 < n=22$) for the front end returns.

4. Alternative asset allocation approaches

In this section, we are going to implement two popular methodologies for optimizing portfolios, such as the Principal Component Analysis, PC, and the mean-variance Utility Maximization, UM, methods, to combine strategies. In both cases, we make use of a fixed 60-month rolling window to extract the portfolio weights for each individual strategy, net the long and short positions of the portfolio constituents, invest the 50% in longs and 50% in shorts, and finally, hold the long-short portfolio for one month. In the Appendix C, we explain the steps for obtaining the long-short PC and the long-short UM portfolios.

Unreported results show both PC and UM approaches provide positive mean returns of 0.40% p.a. (t-stat of 0.19) and 5.63% p.a. (t-stat of 3.36) for front-end returns (0.65% p.a., t-stat of 0.87, and 1.88% p.a., t-stat of 1.91, for spreading returns), respectively. As a comparison with the 1/*K* and multi-sort portfolios, these two portfolios obtain an average Sharpe ratio a 2503% and 66% for front-end returns (and 489% and 99% for spreading returns) higher than the Sharpe ratios of PC and UM portfolios, respectively, for a comparable sample from January 1992 until April 2016. Likewise, the 1/*K* and multi-sort portfolios provide a better crash risk protection, in contrast with both PC and UM. Specifically, the 1/*K* and multi-sort portfolios obtain lower downside volatility, lower 99% VaR and higher maximum drawdown than those of PC and UM portfolios. This may be due to the lack of diversification and the estimation risk that both PC and UM suffer compared with the 1/*K* and multi-sort portfolios.

5. Combined risk premia: Cross-sectional results

A. Methodology

While open to debate (see e.g. Daskalaki and Skiadopoulos, 2014), the literature on commodity futures pricing for the most part tends to highlight the usefulness of the term structure and hedging pressure signals at explaining cross-sectional variations in commodity futures returns (Basu and Miffre, 2013; Szymanowska et al., 2014). In the light of the heightened performance of the combined strategies, it remains to be tested whether the combined 1/*K* and multi-score portfolios could serve as priced risk factor in commodity futures markets in place of the single-score risk premia.

Following the methodology of Fama and MacBeth (1973) as applied in Ang *et al.*'s (2009), we estimate for each commodity i the following time-series regression using the D daily observations available in a given month t

$$r_{id} = \alpha_i + \boldsymbol{\beta}_i'\boldsymbol{f_d} + \varepsilon_{id}\,, \quad d = 1, \dots, D \text{ days} \tag{14}$$

where r_{id} is the excess front-end return (or excess spreading return) of the i^{th} commodity on day d; $\boldsymbol{f}_d$ is a single risk premium or a vector of *K* risk premia based either of front-end contracts or on spreads; ε_{id} is an innovation and $(\alpha_i, \boldsymbol{\beta}_i')'$ is an unknown OLS parameter vector.

At the end of each month t we then estimate the following cross-section regression

$$r_{it} = \lambda_{0t} + \boldsymbol{\lambda}_t' \widehat{\boldsymbol{\beta}}_{it} + \upsilon_{it} \text{ , } i = 1,2, \ldots\ldots N \tag{15}$$

where r_{it} is the month t excess front-end return (or excess spreading return) of the i^{th} commodity; $\widehat{\boldsymbol{\beta}}_{it}$ are the betas contemporaneous to the dependent variable as obtained by OLS estimation of equation (1); and υ_{it} is an innovation. The second step yields a sequence of K monthly vectors of prices of risk, $\boldsymbol{\lambda}_t$, whose statistical significance is tested using Shanken's (1992) error-in-variables consistent standard errors.

The methodology detailed above is applied to either front-end contracts or spreads. The hypothesis tested is whether the front-end risk premia of Table 1 and Table 3, Panel A explain the cross-sectional variations in front-end returns. Likewise, we test whether the spreading risk premia of Table 2 and Table 3, Panel B explain the cross-sectional variations in spreading returns.

B. Empirical results

Table 8 reports average values (and associated *t*-statistics) for the risk premia obtained when each risk factor is treated in isolation. Table 9 reports similar information when more than one risk factor enters the pricing relationships (14) and (15). In total, we consider 14 risk premia as possible sources of priced risk; these includes the equally-weighted long-only portfolio (*EW*), the 11 single-score strategies, and the two combined strategies (1/*K* and multi-score).

<< Insert Tables 8 and 9 around here >>

Table 8, Panel A shows that investors price the risks associated with the 1/*K* and multi-score risk factors. The estimated prices of risk are positive at the 5% level or better, suggesting that investors demand higher expected returns on commodities that load positively on the 1/*K* and multi-score risk premia. Given the significance of the estimated prices of risk, the 1/ *K* and multi-score front-end risk premia are found to successfully explain the cross-section of front-end returns and the 1/*K* and multi-score spreading risk premia are found to successfully explain the cross-section of spreading returns. Table 8, Panel B reports the prices of risk obtained when traditional long-only and single-score premia are used as sources of priced risk, in place of the 1/*K* and multi-score portfolios. With the exception of the term structure signal that explains both front-end and spreading returns, the other stand-alone signals fail to explain both types of returns. All in all, we conclude that the 1/*K* and multi-score portfolios do a better job than any of the risk premia previously identified at explaining cross-sectional futures returns.

These conclusions remain unchanged in Table 9 when the 1/*K* and multi-score portfolios are treated as independent variables in multivariate models that include all risk premia (*EW* and the 11 single-sorts). Based on the results reported in the last set of columns of Table 9, we find as significant factors for front end returns: 1/*K*, multi-sort as well as term structure, speculators' hedging pressure, momentum and skewness. Aside from signals based on 1/*K* and multi-sort, spreading returns are also found to increase with the slope of the term structure, the value of the commodity and the skewness of its returns.

6. Conclusions

Recent research in commodity futures pricing has established that long-short strategies based on individual signals (such as the slope of the term structure of commodity futures prices, past performance, or hedging pressure…) offer a better performance than long-only portfolios. However, instead of devoting efforts to comparing the existing *K* individual signals with a view to helping investors to choose ex-ante one signal, our research aims at unifying the literature by combining the strengths of all *K* signals. This paper proposes long-short combined strategies that combine the individual signals. Our results suggest that the combined signals generate more accurate buy/sell recommendations than either one of the *K* individual signals taken in isolation. The combined signals portfolios are also found to better explain the cross-sectional variations in commodity futures returns better than any of the alternatives previously proposed. All in all, we conclude that the combination strategies we propose are found to capture the risk premium of commodity futures better than individual signals.

References

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X., 2009, High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, 91, 1–23.

Amihud, Y., Mendelson, H., and Lauterbach, B., 1997, Market microstructure and securities values: Evidence from the Tel Aviv Stock Exchange, *Journal of Financial Economics* 45, 365–390.

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

Bakshi, G., Gao, X., and Rossi, A., 2015, Understanding the sources of risk underlying the cross-section of commodity returns, Working Paper.

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.

Bessembinder, H., 1992, Systematic risk, hedging pressure, and risk premiums in futures markets, *Review of Financial Studies* 5, 637-667.

Bodie, Z., and Rosansky, V., 1980, Risk and returns in commodity futures, *Financial Analysts Journal* May/June, 27-39.

Boons, M., and Prado, M.P., 2015, Basis-momentum, Working Paper.

Brennan, M., 1958, The supply of storage, *American Economic Review* 48, 50-72.

Chang, E., 1985, Return to speculators and the theory of normal backwardation, *Journal of Finance* 40, 193-208.

Cootner, P., 1960, Returns to speculators: Telser vs. Keynes, *Journal of Political Economy* 68, 396–404.

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

DeBondt, W. F.M., and Thaler, R., 1985, Does the stock market overreact? *Journal of Finance* 40, 793–805.

DeMiguel, V., Garlappi, L., and Uppal, R., 2007, Optimal Versus Naive Diversification: How Inefficient is the 1/*N* Portfolio Strategy? *Review of Financial Studies*, 22, 1915-1953.

de Roon, F. A., Nijman, T. E., Veld, C., 2000, Hedging pressure effects in futures markets, *Journal of Finance* 55, 1437-1456.

Dhume, D., 2011, Using durable consumption risk to explain commodities returns, Working paper.

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., 1987, Commodity futures prices: some evidence on forecast power, premiums, and the theory of storage, *Journal of Business* 60, 55-73.

Fama, E. F., and MacBeth, J. D., 1973, Risk, returns, and equilibrium: empirical tests, *Journal of Political Economy* 81, 607-636.

Fernandez-Perez, A., Frijns, B., Fuertes, A.-M., and Miffre, J., 2016, The pricing of skewness in commodity futures markets: Risk or lottery?, Working Paper.

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

Gorton, G., Hayashi, F., and Rouwenhorst, G., 2012, The fundamentals of commodity futures returns, *Review of Finance* 17, 35-105.

Hansen, P.R., 2005, A test for superior predictive ability, *Journal of Business and Economic Statistics* 23, 365-380.

Hirshleifer, D., 1988, Residual risk, trading costs, and commodity futures risk premia, *Review of Financial Studies* 1, 173-193.

Hong, J., and Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics* 150, 473-490.

Jegadeesh, N., and Titman, S., 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* 56, 699-720.

Kaldor, N., 1939, Speculation and economic stability, *Review of Economic Studies* 7, 1-27.

Keynes, M., 1930, A Treatise on Money, II: The Applied Theory of Money, edition. Macmillan and Co.

Marshall, Ben R., Nhut H. Nguyen, and Nuttawat Visaltanachoti, 2012, Commodity liquidity measurement and transaction costs, *Review of Financial Studies* 25, 599–638.

Miffre, J., 2016, Long-Short Commodity Investing: A Review of the Literature, *Journal of Commodity Markets* forthcoming.

Miffre, J., and Rallis, G., 2007, Momentum strategies in commodity futures markets, *Journal of Banking and Finance* 31, 6, 1863-1886.

Novy-Marx, R., 2016, Testing strategies based on multiple signals, Working paper.

Locke, P., and Venkatesh, P., 1997, Futures market transaction costs, *Journal of Futures Markets* 17, 229-245.

Opdyke, J.D., 2007, Comparing Sharpe ratios: so, where are the p-values? *Journal of Asset Management* 8, 308–336

Politis, D.N., Romano, J.P., 1994, The stationary bootstrap, *Journal of the American Statistical Association* 89, 1303–1313.

Shanken, J., 1992, On the estimation of beta-pricingmodels, *Review of Financial Studies* 5, 1–33.

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

White, H., 2000, A reality check for data snooping, *Econometrica* 68, 1097–1126.

Working, H., 1949, The theory of price of storage, *American Economic Review* 39, 1254-1262.

Appendix A. Spreading strategies relative to second and third nearest contracts

		Single-sort risk premia											Combined risk premia	
	EW	Term structure	Hedgers' HP	Speculators' HP	Momentum	Value	Volatility	Open interest	Liquidity	FX	Inflation shocks	Skewness	1/K	Multi-sort
Panel A: Spreading returns measured relative to second nearest contracts														
Mean	0.0020	**0.0110**	-0.0013	0.0008	-0.0057	0.0061	-0.0038	-0.0011	0.0016	0.0055	**0.0102**	0.0074	0.0070	**0.0104**
	(0.50)	(2.00)	(-0.30)	(0.16)	(-1.05)	(1.21)	(-0.86)	(-0.26)	(0.36)	(1.17)	(2.20)	(1.61)	(1.55)	(2.12)
StDev	0.0203	0.0309	0.0226	0.0246	0.0300	0.0272	0.0273	0.0245	0.0224	0.0257	0.0268	0.0258	0.0225	0.0246
Downside volatility (0%)	0.0139	0.0182	0.0143	0.0155	0.0196	0.0219	0.0220	0.0201	0.0136	0.0175	0.0151	0.0196	0.0147	0.0137
Skewness	-0.3780	0.5145	0.1776	0.1049	**0.4844**	**-1.0100**	**-1.0038**	**-1.1785**	**0.4398**	-0.0263	**0.3593**	**-0.5697**	-0.0919	**0.4843**
Excess Kurtosis	**3.2051**	**2.3721**	**1.4448**	**1.0278**	**3.9840**	**7.2527**	**6.0749**	**10.0611**	**2.1831**	**1.7233**	**1.8252**	**7.8009**	**1.8808**	**1.3667**
99% VaR (Cornish-Fisher)	0.0191	0.0205	0.0166	0.0176	0.0249	0.0339	0.0326	0.0356	0.0156	0.0200	0.0180	0.0324	0.0178	0.0147
% of positive months	48.58%	53.13%	48.58%	47.73%	46.59%	54.83%	48.86%	52.56%	50.00%	53.13%	51.14%	53.69%	51.99%	54.26%
Maximum drawdown	-0.1577	-0.0801	-0.1699	-0.1276	-0.2231	-0.0949	-0.1641	-0.1168	-0.1283	-0.0945	-0.0573	-0.0955	-0.0815	-0.0911
Sharpe ratio	0.0991	0.3562	-0.0591	0.0311	-0.1911	0.2252	-0.1409	-0.0429	0.0705	0.2148	0.3801	0.2886	0.3120	0.4216
Sortino ratio (0%)	0.1445	0.6059	-0.0932	0.0495	-0.2924	0.2798	-0.1748	-0.0524	0.1164	0.3156	0.6734	0.3795	0.4768	0.7552
Omega ratio	1.0796	1.3215	0.9561	1.0238	0.8569	1.2010	0.8904	0.9647	1.0572	1.1846	1.3333	1.2761	1.2739	1.3897
Panel B: Spreading returns measured relative to 3rd nearest contracts														
Mean	-0.0077	**0.0228**	0.0019	0.0053	-0.0028	**0.0163**	-0.0049	0.0029	-0.0003	0.0042	**0.0140**	**0.0158**	**0.0179**	**0.0215**
	(-1.27)	(2.99)	(0.27)	(0.74)	(-0.39)	(2.29)	(-0.77)	(0.47)	(-0.05)	(0.59)	(2.08)	(2.37)	(2.89)	(3.18)
StDev	0.0287	0.0440	0.0327	0.0352	0.0420	0.0387	0.0375	0.0353	0.0318	0.0372	0.0390	0.0343	0.0313	0.0350
Downside volatility (0%)	0.0194	0.0251	0.0204	0.0210	0.0271	0.0309	0.0275	0.0251	0.0205	0.0258	0.0233	0.0253	0.0191	0.0203
Skewness	**-0.2227**	**0.4948**	**0.2288**	**0.2384**	**0.3590**	**-0.9089**	**-0.2587**	-0.0531	0.1667	-0.0028	0.1720	**-0.4329**	0.0709	**0.3318**
Excess Kurtosis	**1.5284**	**2.2306**	**1.4051**	**0.8624**	**2.5990**	**6.3073**	**4.7262**	**6.0737**	**2.8410**	**2.5186**	**1.3798**	**3.1508**	**0.8890**	**1.6307**
99% VaR (Cornish-Fisher)	0.0241	0.0285	0.0231	0.0233	0.0320	0.0451	0.0393	0.0383	0.0263	0.0310	0.0271	0.0315	0.0209	0.0227
% of positive months	45.45%	53.98%	51.14%	50.85%	47.44%	58.24%	47.16%	51.70%	47.73%	51.70%	54.55%	58.24%	55.68%	54.55%
Maximum drawdown	-0.3705	-0.0972	-0.2455	-0.2022	-0.2092	-0.1025	-0.2575	-0.1435	-0.1979	-0.1910	-0.1279	-0.1807	-0.1047	-0.0940
Sharpe ratio	-0.2671	0.5190	0.0579	0.1501	-0.0670	0.4224	-0.1313	0.0821	-0.0090	0.1138	0.3595	0.4606	0.5723	0.6157
Sortino ratio (0%)	-0.3943	0.9091	0.0926	0.2520	-0.1036	0.5283	-0.1789	0.1155	-0.0140	0.1643	0.6017	0.6253	0.9342	1.0632
Omega ratio	0.8169	1.4919	1.0451	1.1195	0.9476	1.4060	0.8969	1.0693	0.9929	1.0951	1.3112	1.4376	1.5491	1.6121

Appendix B. Spreading strategies (1st - 4th contracts) using the holding returns as in Szymanowska *et al.*, 2014

		Single-sort risk premia											**Combined risk premia**	
	EW	Term structure	Hedgers' HP	Speculators' HP	Momentum	Value	Volatility	Open interest	Liquidity	FX	Inflation shocks	Skewness	1/K	Multi-sort
Mean	-0.0073	**0.0143**	0.0021	0.0055	0.0004	**0.0148**	-0.0089	-0.0003	-0.0046	0.0029	0.0071	**0.0118**	**0.0107**	**0.0153**
	(-1.54)	(2.07)	(0.34)	(0.86)	(0.06)	(2.18)	(-1.70)	(-0.06)	(-0.87)	(0.45)	(1.25)	(2.05)	(1.85)	(2.50)
StDev	0.0230	0.0375	0.0299	0.0311	0.0367	0.0349	0.0313	0.0316	0.0280	0.0326	0.0354	0.0320	0.0289	0.0319
Downside volatility (0%)	0.0168	0.0233	0.0189	0.0194	0.0236	0.0266	0.0245	0.0230	0.0191	0.0260	0.0234	0.0254	0.0192	0.0197
Skewness	**-0.4815**	**0.2691**	**0.3402**	0.0657	**0.4242**	**-0.6391**	**-0.8754**	**-0.3962**	**0.2402**	**-0.3411**	0.1158	**-0.8330**	-0.0545	**0.4745**
Excess Kurtosis	**3.2783**	**1.7181**	**2.5186**	**0.7933**	**2.8666**	**4.4437**	**5.7396**	**5.1220**	**3.6191**	**4.9271**	**4.2882**	**6.7626**	**1.6938**	**3.4800**
99% VaR (Cornish-Fisher)	0.0229	0.0259	0.0224	0.0217	0.0277	0.0358	0.0371	0.0342	0.0244	0.0344	0.0325	0.0384	0.0221	0.0237
% of positive months	45.17%	56.53%	50.85%	53.13%	51.99%	54.83%	44.89%	46.31%	48.01%	53.69%	50.00%	57.95%	53.98%	54.83%
Maximum drawdown	-0.2598	-0.1140	-0.2096	-0.2122	-0.1298	-0.1049	-0.2647	-0.1642	-0.2386	-0.2121	-0.1119	-0.1657	-0.1395	-0.1185
Sharpe ratio	-0.3168	0.3807	0.0710	0.1753	0.0106	0.4242	-0.2854	-0.0104	-0.1652	0.0875	0.2020	0.3694	0.3694	0.4801
Sortino ratio (0%)	-0.4343	0.6126	0.1123	0.2819	0.0165	0.5554	-0.3651	-0.0143	-0.2423	0.1097	0.3055	0.4657	0.5560	0.7779
Omega ratio	0.7814	1.3441	1.0574	1.1417	1.0086	1.4173	0.7896	0.9913	0.8742	1.0777	1.1758	1.3536	1.3378	1.4760

Appendix C. Alternative asset allocation approaches

In this Appendix, we explain the steps followed to obtain the weights for the Principal Component (PC) and Utility Maximization (UM) portfolios. For simplicity, we focus on the front end returns, though the steps are the same for the spreading returns.

Principal Component Analysis:

We employ a rolling window of 60 months to extract the first principal component of the monthly returns for the 11 individual signals.[9] Using the loadings of each individual strategy on the first principal component, we extract the weights for each individual strategy as follows,

$$w_{i,t}^{PC} = \frac{c_{i,t}^{PC}}{\sum_{i=1}^{K}\left|c_{i,t}^{PC}\right|} \qquad (C.1)$$

where $c_{i,t}^{PC}$ is the loading of the strategy *i*-th on the first principal component over a 60-month window ending and including month t. We then net the individual positions and ensure that long and short recommendations for a given commodity cancel out in the PC portfolio. Finally, we hold the net positions on a fully-collateralized basis over the following month ensuring that at month-end 50% of the client's mandate is invested in the longs and 50% is invested in the shorts. All these steps are repeated until the end of the sample.

Utility maximization:

We estimate the weights of this strategy maximizing the utility function of a mean-variance investor with a Relative Risk Aversion (RRA) of 5 over a rolling window of 60 months, employing the monthly returns of the 11 individual signals. Specifically, we optimize the mean-variance utility function, imposing that all the weights are positive and the entire budget is invested. This would be in matrix notation,

[9] On average, the 1st PC is 32.50% for front end returns and 33.98% for spreading returns.

$$
\underset{\vec{\mathrm{w}}_{\mathrm{t}}^{\mathrm{UM}}}{\mathrm{Min}}\left\{\frac{\mathrm{RRA}}{2}\cdot\vec{\mathrm{w}}_{\mathrm{t}}^{\mathrm{UM}'}\cdot\Sigma_t\cdot\vec{\mathrm{w}}_{\mathrm{t}}^{\mathrm{UM}}-\vec{\mathrm{w}}_{\mathrm{t}}^{\mathrm{UM}'}\cdot\mu_t\right\}
$$

$s.t.$

$$
\sum_{i=1}^{K} w_{i,t}^{UM}=1
$$

$$
\forall w_{i,t}^{UM}\geq 0; i=1,\dots,K
$$

(C.2)

where Σ_t is the variance-covariance matrix of the individual strategies over a 60-month window ending and including month t, and μ_t is the a vector of mean returns for the individual strategies over a 60-month window ending and including month t. We then net the individual positions and ensure that long and short recommendations for a given commodity cancel out in the UM portfolio. Finally, we hold the net positions on a fully-collateralized basis over the following month ensuring that at month-end 50% of the client's mandate is invested in the longs and 50% is invested in the shorts. All these steps are repeated until the end of the sample.

Table 1. Front-end equally-weighted and single-score strategies

	EW	Term structure	Hedgers' HP	Speculators' HP	Momentum	Value	Volatility	Open interest	Liquidity	FX	Inflation shocks	Skewness
Panel A: Summary statistics												
Mean	-0.0077	**0.0677**	0.0255	**0.0464**	**0.0621**	0.0240	0.0102	0.0060	-0.0020	0.0114	0.0311	**0.0522**
	(-0.28)	(3.62)	(1.31)	(2.27)	(3.08)	(1.14)	(0.58)	(0.35)	(-0.10)	(0.58)	(1.44)	(2.60)
StDev	0.1294	0.1034	0.0923	0.0998	0.1156	0.1168	0.0952	0.0943	0.0904	0.1081	0.1184	0.0967
Downside volatility (0%)	0.0972	0.0639	0.0540	0.0616	0.0741	0.0752	0.0640	0.0613	0.0570	0.0803	0.0715	0.0600
Skewness	**-0.6895**	0.0129	0.2135	**0.4906**	-0.0676	-0.0950	**-0.1844**	0.0259	0.0357	**-0.5226**	**0.2632**	0.0009
Excess Kurtosis	**4.1211**	**1.9296**	**0.5940**	**4.1753**	**0.8368**	**0.6013**	**1.7894**	**0.9566**	**0.5302**	**1.3998**	**1.6803**	**0.8490**
99% VaR (Cornish-Fisher)	0.1358	0.0770	0.0589	0.0782	0.0806	0.0834	0.0780	0.0684	0.0634	0.0906	0.0829	0.0661
% of positive months	51.99%	57.95%	54.26%	58.81%	55.40%	51.70%	51.42%	51.42%	50.85%	53.13%	51.42%	56.53%
Maximum drawdown	-0.6029	-0.2207	-0.4906	-0.4792	-0.3049	-0.5018	-0.3260	-0.4615	-0.4750	-0.3743	-0.3135	-0.3743
Sharpe ratio	-0.0592	**0.6543**	0.2760	**0.4649**	**0.5373**	0.2058	0.1066	0.0636	-0.0218	0.1051	0.2627	**0.5397**
Sortino ratio (0%)	-0.0788	1.0591	0.4713	0.7525	0.8384	0.3200	0.1586	0.0978	-0.0346	0.1415	0.4354	0.8703
Omega ratio	0.9546	**1.6554**	1.2285	**1.4304**	**1.5111**	1.1661	1.0844	1.0495	0.9838	1.0836	1.2221	**1.4990**
Panel B: Correlation matrix												
Term structure	**0.09**											
Hedgers' HP	**0.11**	0.00										
Speculators' HP	0.05	**0.11**	**0.73**									
Momentum	**0.12**	**0.40**	**0.14**	**0.33**								
Value	**-0.18**	**-0.35**	**-0.15**	**-0.26**	**-0.43**							
Volatility	**0.21**	0.06	0.02	0.03	0.04	**-0.22**						
Open interest	0.02	**-0.09**	0.03	0.00	**-0.18**	0.07	0.02					
Liquidity	0.02	**0.11**	**-0.11**	**-0.18**	-0.04	-0.01	0.05	0.02				
FX	**0.25**	0.02	-0.02	0.06	-0.06	**-0.12**	**0.16**	**-0.13**	**0.12**			
Inflation shocks	**0.16**	**0.24**	-0.06	**-0.17**	**0.13**	**-0.23**	-0.02	-0.07	**0.31**	0.03		
Skewness	0.04	0.08	**0.16**	**0.21**	-0.02	**-0.12**	0.05	-0.03	-0.03	0.05	0.07	

Panel A presents summary statistics for long-only and long-short risk premia where the latter are based on an unique signal. Panel B presents cross risk premia correlations. The strategies are implemented on front-end returns thereby measuring spot premia. EW stands for a long-only equally-weighted monthly-rebalanced portfolio of all commodity futures. HP stands for hedging pressure. Mean and standard deviation (StDev) are annualized. Newey-West *t*-statistics are in parentheses. Bold signifies significance at the 10% level or better. The sample covers the period January 1987-April 2016.

Table 2. Spreading equally-weighted and single-score strategies

	EW	Term structure	Hedgers' HP	Speculators' HP	Momentum	Value	Volatility	Open interest	Liquidity	FX	Inflation shocks	Skewness
Panel A: Summary statistics												
Mean	**-0.0147**	**0.0295**	0.0085	0.0117	0.0014	**0.0194**	0.0002	0.0021	-0.0010	0.0032	**0.0137**	**0.0214**
	(-2.01)	(3.37)	(0.97)	(1.32)	(0.16)	(2.13)	(0.02)	(0.29)	(-0.13)	(0.37)	(1.68)	(2.54)
StDev	0.0347	0.0518	0.0407	0.0427	0.0510	0.0477	0.0438	0.0422	0.0385	0.0454	0.0480	0.0414
Downside volatility (0%)	0.0227	0.0304	0.0250	0.0255	0.0319	0.0367	0.0296	0.0295	0.0234	0.0328	0.0298	0.0298
Skewness	-0.0708	**0.2869**	**0.2388**	0.2027	**0.3136**	**-0.6570**	**0.0215**	-0.0021	**0.2792**	**-0.2263**	0.0326	**-0.3848**
Excess Kurtosis	**0.4578**	**1.9903**	**1.5503**	**0.9728**	**2.1904**	**4.9818**	**4.0215**	**4.7938**	**2.2105**	**3.0522**	**1.2045**	**2.5725**
99% VaR (Cornish-Fisher)	0.0261	0.0357	0.0286	0.0285	0.0377	0.0509	0.0411	0.0418	0.0291	0.0415	0.0347	0.0359
% of positive months	44.89%	54.55%	53.41%	51.99%	48.58%	57.67%	48.01%	51.70%	47.16%	54.26%	55.11%	57.95%
Maximum drawdown	-0.4763	-0.1030	-0.2691	-0.2319	-0.1843	-0.1263	-0.2147	-0.1764	-0.2299	-0.2559	-0.1613	-0.2187
Sharpe ratio	-0.4242	0.5699	0.2078	0.2736	0.0284	0.4076	0.0043	0.0496	-0.0265	0.0703	0.2857	0.5160
Sortino ratio (0%)	-0.6499	0.9706	0.3383	0.4582	0.0455	0.5303	0.0063	0.0708	-0.0436	0.0976	0.4603	0.7178
Omega ratio	0.7305	1.5420	1.1705	1.2260	1.0226	1.3851	1.0035	1.0407	0.9799	1.0580	1.2416	1.4974
Panel B: Correlation matrix												
Term structure	**0.25**											
Hedgers' HP	**-0.11**	0.06										
Speculators' HP	0.04	**0.20**	**0.65**									
Momentum	**0.17**	**0.48**	**0.18**	**0.35**								
Value	**-0.16**	**-0.44**	**-0.18**	**-0.23**	**-0.52**							
Volatility	**0.24**	0.04	-0.01	-0.02	-0.01	**-0.12**						
Open interest	0.03	0.05	-0.02	-0.02	-0.03	0.05	0.07					
Liquidity	**0.23**	0.03	0.00	**-0.14**	-0.06	0.07	0.03	**-0.11**				
FX	**0.13**	0.05	-0.08	0.02	-0.08	-0.02	0.04	-0.08	0.04			
Inflation shocks	0.05	**0.14**	0.00	-0.03	**0.16**	**-0.20**	-0.04	-0.05	**0.27**	**-0.11**		
Skewness	0.05	0.02	**0.09**	**0.22**	-0.08	0.00	**0.10**	0.05	-0.01	**0.20**	-0.05	

Panel A presents summary statistics for long-only and long-short risk premia where the latter are based on an unique signal. Panel B presents cross risk premia correlations. The strategies are implemented on spreading returns thereby measuring the difference between spot and term premia. EW stands for a long-only equally-weighted monthly-rebalanced portfolio of all commodity futures. HP stands for hedging pressure. Mean and standard deviation (StDev) are annualized. Newey-West *t*-statistics are in parentheses. Bold signifies significance at the 10% level or better. The sample covers the period January 1987-April 2016.

Table 3. 1/*K* versus multi-score risk premia

	1/K					Multi-sort					Best single-sort
	Q1	Q2	Q3	Q4	Q4-Q1	Q1	Q2	Q3	Q4	Q4-Q1	
Panel A: Front-end strategies											
Mean	**-0.0784**	-0.0328	0.0156	**0.0804**	**0.0794**	**-0.0873**	-0.0239	-0.0141	**0.0913**	**0.0893**	**0.0677**
	(-2.99)	(-1.09)	(0.48)	(2.16)	(5.29)	(-3.11)	(-0.75)	(-0.42)	(2.31)	(5.03)	(3.62)
StDev	0.1395	0.1612	0.1576	0.1823	0.0829	0.1475	0.1683	0.1714	0.1865	0.0924	0.1034
Downside volatility (0%)	0.0962	0.1204	0.1074	0.1310	0.0454	0.0961	0.1217	0.1223	0.1273	0.0504	0.0639
Skewness	-0.1734	**-0.7280**	-0.2117	**-0.7261**	0.0975	-0.0072	**-0.6041**	**-0.2219**	**-0.4410**	0.1978	0.0129
Excess Kurtosis	**0.8860**	**4.4407**	**1.1496**	**4.1358**	0.2586	**0.6650**	**3.9836**	**1.7146**	**2.2783**	0.2298	**1.9296**
99% VaR (Cornish-Fisher)	0.1133	0.1750	0.1231	0.1843	0.0487	0.1132	0.1752	0.1432	0.1599	0.0518	0.0770
% of positive months	43.75%	49.43%	50.85%	55.68%	58.52%	43.18%	48.01%	50.85%	56.53%	59.94%	57.95%
Maximum drawdown	-0.9372	-0.8542	-0.5063	-0.5465	-0.1590	-0.9597	-0.8043	-0.7543	-0.5400	-0.1604	-0.2207
Sharpe ratio	-0.5616	-0.2037	0.0990	0.4408	0.9578	-0.5923	-0.1422	-0.0824	0.4895	0.9665	0.6543
Sortino ratio (0%)	-0.8146	-0.2728	0.1453	0.6134	1.7481	-0.9092	-0.1966	-0.1154	0.7175	1.7718	1.0591
Omega ratio	0.6539	0.8525	1.0791	1.3974	2.0363	0.6450	0.8949	0.9375	1.4441	2.0778	1.6554
Panel B: Spreading strategies											
Mean	**-0.0392**	**-0.0198**	-0.0138	0.0123	**0.0257**	**-0.0411**	-0.0012	**-0.0347**	0.0184	**0.0298**	**0.0295**
	(-3.93)	(-2.09)	(-1.27)	(0.92)	(3.52)	(-4.21)	(-0.11)	(-2.82)	(1.28)	(3.71)	(3.37)
StDev	0.0461	0.0520	0.0524	0.0669	0.0378	0.0507	0.0545	0.0614	0.0727	0.0433	0.0518
Downside volatility (0%)	0.0358	0.0360	0.0385	0.0435	0.0229	0.0393	0.0314	0.0483	0.0481	0.0261	0.0304
Skewness	-0.6830	-0.1408	-0.5306	0.1276	-0.0075	-0.6793	**0.5499**	-1.1150	0.1193	0.1854	0.2869
Excess Kurtosis	**2.5970**	**2.6085**	**2.8174**	**2.1874**	**0.9477**	**3.9326**	**1.5947**	**7.5485**	**5.1333**	**2.6771**	**1.9903**
99% VaR (Cornish-Fisher)	0.0466	0.0471	0.0506	0.0519	0.0257	0.0557	0.0344	0.0816	0.0706	0.0325	0.0357
% of positive months	38.92%	43.47%	46.31%	50.00%	58.81%	39.77%	47.73%	44.89%	51.14%	55.11%	54.55%
Maximum drawdown	-0.7403	-0.5883	-0.5153	-0.4262	-0.1354	-0.7633	-0.3283	-0.6881	-0.3978	-0.1158	-0.1030
Sharpe ratio	-0.8500	-0.3819	-0.2627	0.1831	0.6796	-0.8119	-0.0226	-0.5658	0.2535	0.6886	0.5699
Sortino ratio (0%)	-1.0945	-0.5507	-0.3577	0.2813	1.1234	-1.0460	-0.0393	-0.7187	0.3833	1.1419	0.9706
Omega ratio	0.5097	0.7411	0.8143	1.1591	1.6681	0.5189	0.9826	0.6351	1.2319	1.7090	1.5420

The table presents summary statistics for the combined strategies based on either 1/*K* or multiple scores (details on the strategies can be found in Section 3.A). Q1,…, Q4 are quartiles portfolios based on the combined signals and Q4-Q1 measures the combined risk premium. The strategies are implemented on front-end contracts in Panel A thereby measuring spot premia. The strategies are implemented on spreads in Panel B thereby measuring the difference between spot and term premia. The last column pertains to the front-end and spreading strategies that have the highest Sharpe ratios in Tables 1 and 2; this information is presented here to facilitate comparison. Mean and standard deviation (StDev) are annualized. Newey-West *t*-statistics are in parentheses. Bold signifies significance at the 10% level or better. The sample covers the period January 1987-April 2016.

Table 4. Risk and performance

	Front-end risk premia				Spreading risk premia			
	1/*K*		**Multi-score**		**1/*K***		**Multi-score**	
Alpha	**0.0040**	(1.83)	0.0112	(1.41)	0.0005	(0.41)	0.0037	(0.89)
EW	0.0118	(1.58)	-0.0087	(-0.40)	**0.0360**	(3.15)	-0.0372	(-0.69)
Term structure	**0.2173**	(22.74)	**0.2260**	(9.43)	**0.2274**	(22.40)	**0.2653**	(6.45)
Hedgers' HP	**0.2183**	(16.77)	**0.1933**	(3.06)	**0.2383**	(18.34)	**0.2412**	(5.08)
Speculators' HP	**0.2649**	(19.72)	**0.2592**	(4.61)	**0.2568**	(16.56)	**0.2143**	(3.42)
Momentum	**0.2092**	(18.33)	**0.2387**	(7.79)	**0.2181**	(19.71)	**0.2374**	(6.38)
Value	**0.2301**	(24.25)	**0.2087**	(6.75)	**0.2526**	(25.07)	**0.1828**	(3.34)
Volatility	**0.2229**	(20.67)	**0.1698**	(5.47)	**0.2285**	(22.11)	**0.1464**	(3.09)
Open interest	**0.2472**	(21.78)	**0.2611**	(8.99)	**0.2399**	(21.82)	**0.2201**	(5.72)
Liquidity	**0.2457**	(23.26)	**0.2368**	(7.12)	**0.2590**	(25.15)	**0.2194**	(3.94)
FX	**0.2079**	(15.75)	**0.2394**	(8.34)	**0.2025**	(16.69)	**0.2404**	(5.14)
Inflation shocks	**0.2297**	(22.74)	**0.2246**	(7.68)	**0.2097**	(17.00)	**0.2145**	(4.60)
Skewness	**0.2224**	(18.87)	**0.2456**	(6.70)	**0.2343**	(19.69)	**0.2472**	(6.07)
Adjusted R^2	97.03%		77.87%		96.90%		72.03%	

Table 4 report estimated coefficients and *t*-statistics (corrected by Newey-West standard errors) obtained by regressing front-end (spreading) combined returns on the front-end (spreading) risk factors of Table 1 (Table 2). EW stands for a long-only equally-weighted monthly-rebalanced portfolio of all commodity futures. HP stands for hedging pressure. Bold signifies significance at the 10% level or better. The sample covers the period January 1987-April 2016.

Table 5. Transaction costs and net performance

	Front-end risk premia		Spreading risk premia	
	1/*K*	**Multi-score**	**1/*K***	**Multi-score**
Panel A: Net performance				
		T-costs = 0.033%		
Mean	**0.0774**	**0.0865**	**0.0219**	**0.0252**
	(5.15)	(4.86)	(3.00)	(3.12)
Sharpe ratio	0.9336	0.9332	0.5804	0.5798
		T-costs = 0.066%		
Mean	**0.0753**	**0.0842**	**0.0182**	**0.0206**
	(5.02)	(4.73)	(2.49)	(2.55)
Sharpe ratio	0.9093	0.9089	0.4811	0.4733
Panel B: Break-even analysis and turnover				
Break-even T-costs (%)	1.303	1.303	0.456	0.436
Turnover	0.506	0.600	0.945	1.172

Panel A presents summary statistics for the combined 1/K and multi-score strategies after accounting for a round-trip transaction cost of 0.033% and 0.066%. Panel B presents a break-even analysis and turnover. Newey-West *t*-statistics are in parentheses. Bold signifies significance at the 10% level or better. The sample covers the period January 1987-April 2016.

Table 6. Data Snooping

		Hansen' SPA test			
		Front-end risk premia		***Spreading risk premia***	
Benchmark	*Loss function*	*Consistent p-values*		*Consistent p-values*	
		q=0.2	*q=0.5*	*q=0.2*	*q=0.5*
1/K	Linear	0.3955	0.5006	0.4283	0.4510
	Exp(λ = 1)	0.4662	0.5509	0.4325	0.4497
	Exp(λ = 2)	0.5160	0.6141	0.4553	0.4582
Multi-score	Linear	0.7741	0.8365	0.2746	0.3061
	Exp(λ = 1)	0.7883	0.8475	0.2696	0.2988
	Exp(λ = 2)	0.8001	0.8503	0.2717	0.2967
EW-long only	Linear	0.0033	0.0011	0.0000	0.0000
	Exp(λ = 1)	0.0029	0.0014	0.0000	0.0000
	Exp(λ = 2)	0.0013	0.0011	0.0000	0.0000

P-values for the Hansen's (2005) SPA test using three loss function and two q-values for 1/K, multi-score and EW-long only as benchmarks. The sample covers the period January 1987-April 2016.

Table 7. Novy-Marx's (2016) Selection and Overfitting test.

	Front-end risk premia		Spreading risk premia	
	1/*K*	**Multi-score**	**1/*K***	**Multi-score**
Panel A: Pure overfitting (*K = n = 11*)				
Novy-Marx statistic	5.22	6.50	3.90	5.32
Critical values (1%)	*4.07*	*4.97*	*4.07*	*4.97*
Critical values (5%)	*3.66*	*4.44*	*3.66*	*4.44*
Critical values (10%)	*3.45*	*4.16*	*3.45*	*4.16*
Panel B: Selection and overfitting biases (1 < *K=11* < *n*)				
n for rejection at 1% level	18	35	11	17
n for rejection at 5% level	23	44	13	21
n for rejection at 10% level	26	51	14	23

Panel A reports the Novy-Marx's (2016) pure overfitting test for 1/K and multi-score, and their critical values at the {1%, 5%, 10%} level. Panel B reports the number of true candidate signals, *n*, that makes the performance of the combined strategies statistically insignificant at the {1%, 5%, 10%} level. The sample covers the period January 1987-April 2016.

Table 8. Single-factor cross-sectional analysis

	Front-end risk premia			Spreading risk premia		
	λ	$t(\lambda)$	Adj-R^2	λ	$t(\lambda)$	Adj-R^2
Panel A: Combined risk premia						
1/K	**0.0024**	**(2.12)**	9.15%	**0.0021**	**(5.85)**	14.66%
Multi-score	**0.0043**	**(4.01)**	9.57%	**0.0024**	**(6.03)**	15.25%
Panel B: Traditional risk premia						
EW	0.0027	(1.36)	9.98%	**0.0015**	**(3.37)**	15.06%
Term structure	**0.0024**	**(1.65)**	10.40%	**0.0021**	**(2.98)**	16.96%
Hedgers' HP	0.0002	(0.18)	10.12%	0.0005	(0.75)	14.15%
Speculators' HP	0.0012	(0.86)	10.40%	**0.0012**	**(1.98)**	14.99%
Momentum	0.0024	(1.42)	11.66%	0.0004	(0.54)	14.80%
Value	**0.0036**	**(2.32)**	11.46%	0.0009	(1.26)	13.72%
Volatility	-0.0015	(-1.07)	9.52%	-0.0003	(-0.48)	13.35%
Open interest	0.0018	(1.32)	10.05%	0.0000	(-0.02)	13.72%
Liquidity	0.0006	(0.40)	9.95%	0.0003	(0.50)	13.79%
FX	-0.0013	(-0.78)	11.28%	-0.0004	(-0.50)	14.18%
Inflation shocks	0.0014	(0.80)	11.89%	0.0010	(1.34)	16.31%
Skewness	0.0004	(0.25)	9.84%	**0.0012**	**(1.98)**	13.11%

The table presents averages of the prices of risk λ estimated from second-stage cross-sectional regressions using either front-end contracts or spreads as base assets and either one of the risk premia as independent variables. Shanken-corrected *t*-statistics are in parentheses. Bold signifies significance at the 10% level or better. The sample covers the period January 1987-April 2016.

Table 9. Multiple-factor cross-sectional analysis

	λ	$t(\lambda)$	λ	$t(\lambda)$	λ	$t(\lambda)$
Panel A: Front-end risk premia						
Intercept	-0.0005	(-0.20)	0.0004	(0.17)	-0.0014	(-0.55)
1/K					**0.0060**	**(4.53)**
Multi-score			**0.0069**	**(4.64)**		
EW	0.0025	(0.94)	0.0017	(0.60)	0.0035	(1.27)
Term structure	**0.0047**	**(2.78)**	**0.0050**	**(2.90)**	**0.0052**	**(3.02)**
Hedgers' HP	0.0016	(1.08)	0.0014	(0.90)	0.0017	(1.10)
Speculators' HP	**0.0030**	**(1.80)**	**0.0029**	**(1.77)**	**0.0029**	**(1.77)**
Momentum	**0.0043**	**(2.26)**	**0.0046**	**(2.41)**	**0.0046**	**(2.40)**
Value	0.0024	(1.27)	0.0021	(1.13)	0.0022	(1.16)
Volatility	0.0003	(0.22)	0.0005	(0.29)	0.0005	(0.29)
Open interest	0.0008	(0.53)	0.0009	(0.56)	0.0009	(0.60)
Liquidity	-0.0002	(-0.12)	0.0001	(0.04)	0.0001	(0.07)
FX	0.0007	(0.41)	0.0006	(0.36)	0.0007	(0.42)
Inflation shocks	0.0032	(1.69)	0.0031	(1.60)	0.0031	(1.60)
Skewness	**0.0033**	**(2.07)**	**0.0034**	**(2.13)**	**0.0035**	**(2.23)**
Adj-R^2	54.92%		56.22%		57.64%	
Panel B: Spreading risk premia						
Intercept	**-0.0012**	**(-3.67)**	**-0.0012**	**(-3.37)**	**-0.0011**	**(-3.25)**
1/K					**0.0023**	**(3.55)**
Multi-score			**0.0025**	**(3.39)**		
EW	0.0007	(1.10)	0.0006	(0.94)	0.0006	(0.96)
Term structure	**0.0023**	**(2.66)**	**0.0024**	**(2.72)**	**0.0025**	**(2.76)**
Hedgers' HP	0.0009	(1.31)	0.0009	(1.27)	0.0009	(1.27)
Speculators' HP	0.0011	(1.53)	0.0011	(1.54)	0.0010	(1.41)
Momentum	0.0002	(0.30)	0.0003	(0.32)	0.0001	(0.17)
Value	0.0012	(1.57)	**0.0013**	**(1.71)**	**0.0014**	**(1.85)**
Volatility	0.0000	(0.04)	0.0000	(0.03)	0.0001	(0.16)
Open interest	0.0002	(0.34)	0.0002	(0.30)	0.0003	(0.39)
Liquidity	0.0004	(0.62)	0.0004	(0.54)	0.0004	(0.67)
FX	0.0002	(0.30)	0.0003	(0.37)	0.0001	(0.19)
Inflation shocks	0.0009	(1.07)	0.0009	(1.07)	0.0010	(1.24)
Skewness	**0.0015**	**(2.31)**	**0.0015**	**(2.26)**	**0.0016**	**(2.48)**
Adj-R^2	72.85%		76.16%		76.03%	

The table presents averages of the prices of risk λ estimated from second-stage cross-sectional regressions using either front-end contracts (Panel A) or spreads (Panel B) as base assets and combinations of front-end (Panel A) or spreading (Panel B) risk premia as independent variables. Shanken-corrected *t*-statistics are in parentheses. Bold signifies significance at the 10% level or better. The sample covers the period January 1987-April 2016.

Figure 1. Future Value of $1 Invested in long-only and long-short fully-collateralized front-end portfolios

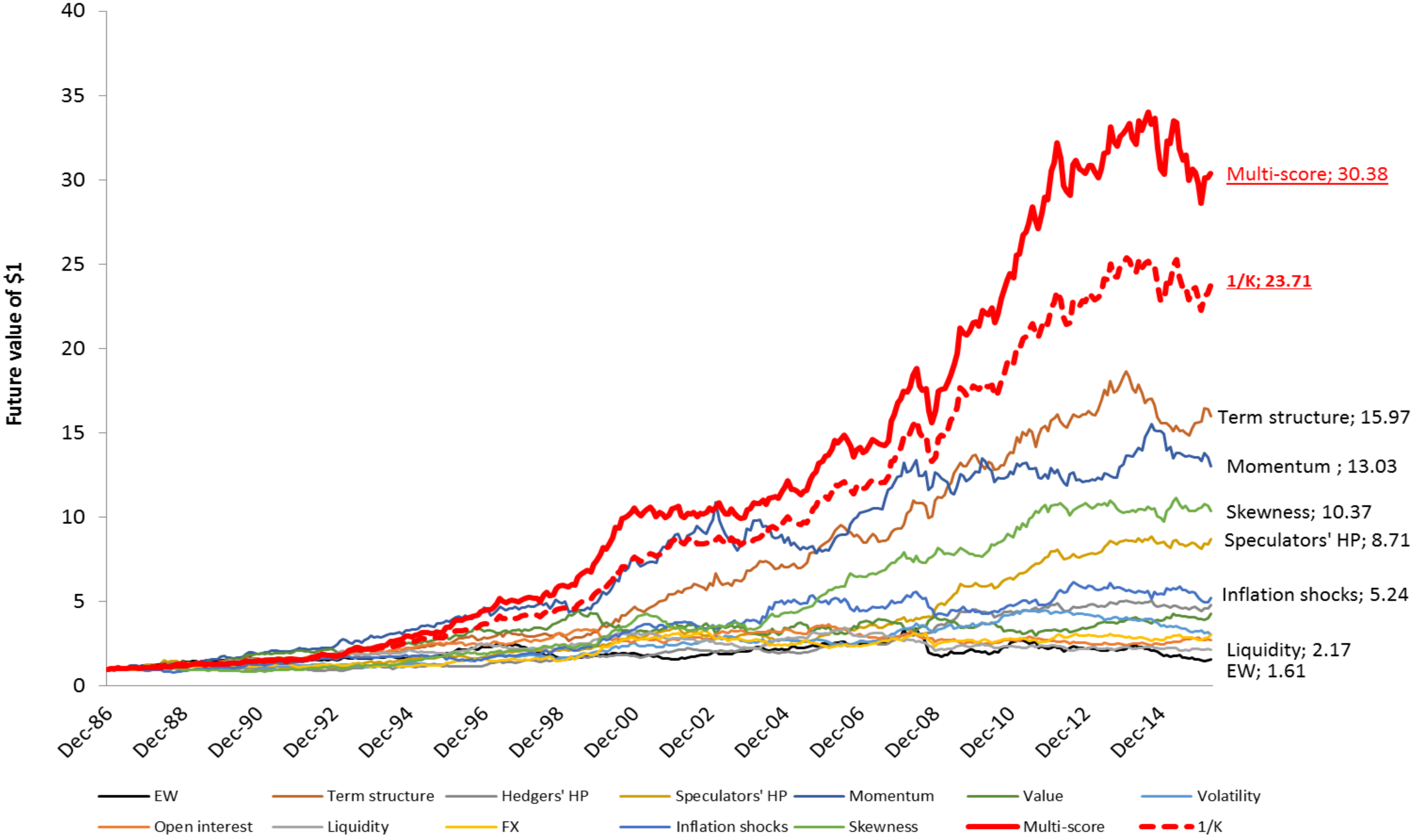

Figure 2. Future Value of $1 Invested in long-only and long-short fully-collateralized spreading portfolios