LUISS GUIDO CARLI

MASTER THESIS

Active Management of Commodities through Signals

Author:
Xavier MAGI

Supervisor:
Prof. Pierpaolo BENIGNO

Co-Supervisor:
Prof. Alberto CYBO-OTTONE

LUISS
1977-2017 Quarantesimo

LUISS Guido Carli

Abstract

Department of Economics and Finance

Master in Finance

Active Management of Commodities through Signals

by Xavier MAGI

Commodity futures contracts have become an important asset class for investors and a source of diversification. This study demonstrates that combining signals based on capturing the fundamental of the slope of term structure can yield Sharpe ratio much higher than the S&P-GSCI index. Furthermore, we implement a neural network based combination strategy to further improve the performance of our portfolio. Robustness tests are performed to further strengthen our claim.

Keywords: Commodity Futures, Forecasting, Signals, Momentum, Term Structure, Idiosyncratic Volatility, Hedging Pressure, Higher Moments, Neural Network.

Contents

1 Introduction

In the last decade, commodity futures have seen an huge increase in their popularity among portfolio investors who seek diversification and improved performance, making commodities an asset class of their own. Although theories on the slope of the commodity future contracts, the longstanding hedging pressure theory, date back to Keynes(1923), Hicks(1939) and Hirshleifer(1988), the literature on how to construct signals to capture risk premium based on the fundamentals of backwardation and contango is quite recent. These signals are constructed from the slope of the term structure (TS) of commodity futures prices (Gorton&Rouwenhorst, 2006; Erb & Harvey, 2006); Idiosyncratic Volatility by Miffre *et al*,2013 (IVol); ; inventory levels (Gorton *et al*, 2012), hedging pressure (Basu & Miffre,2013; Dewally *et al*,2013); Momentum (Mom) (Miffre & Rallis,2007; Szakmary *et al*, 2010; Erb & Harvey, 2006; Shen *et al*,2007) and finally Skewness (SK) (Perez *et al*, 2015). Although the profitability of the last two signal (Mom and SK) have been empirically proven, the theory and hence the reason behind the profits is still under debate. For the momentum strategy, Barberis *et al* (1998), Daniel *et al* (1998), Hong & Stein (1999) argue for behavioral models which attribute abnormal returns to over-and-under reactions of investors to news, while Chordia & Shivakumar (2002) and Lesmond *et al* (2004) attribute them to transaction costs or time-variation in expected returns.

This study concerns the implementation of the Momentum, Term Structure, Idiosyncratic Volatility, Hedging Pressure Signals and Skewness and the added Kurtosis signals. TS and Hedging Pressure signals originate directly from the commodity literature of theory of storage (Working,1949; Brennan, 1958). Mom and IVol arise from the equity pricing literature with Jegadeesh & Titman (1993) proposing a long-short portfolio going long with the recent best performers and going short witht the recent worst performers. The IVol strategy consists in buying stock with low idiosyncratic risk and sell stock with high idiosyncratic risk as proposed by Ang *et al*

(2006,2009). Furthermore, we will show that the signals are non-overlapping, motivating the combination of the studied signals to enhance performance. The various risk-return ratios (Sharpe, Sortino, Omega) confirm that the signals (by themselves and combined) provide better returns in respect to the commodity indexes we consider, namely the Standard & Poor's Goldman Sachs Commodity Index (S&P-GSCI) and the Bloomberg Commodity Index (BCOM). Moreover, we discuss how to combine our signals in order to maximize risk-return. One of the combination model defined will rely on the implementation of Artificial Neural Network (ANN), which a computational method used in machine learning. Even though the theory behind machine learning has been around for a long time, the application of it (especially in finance) is quite recent. Before the actual implementation is presented, we display a simple neural network example, to acquaint readers new to the concept of machine learning to our neural network model.

1.1 The Financialization of Commodity Markets

Figure 1 plot the index level of the S&P-GSCI, BCOM and the equally weighted portfolio constructed with the commodities studied in this paper. Moreover, the below graph plots the average open interest (the sum of the total contract that have not been settled, outstanding) of the studied commodities. The figure shows that after the early 2000s the commodity index levels rose exponentially, closely followed by the open interest. This trend has continued until 2008, year in which the commodity prices fell steeply, mostly caused by the slowdown of the world economy. The large inflow of investment capital (represented by open interest) to commodity futures markets has generated a debate about whether the so-called "financialization" has distorted commodity prices. Cheng & Xiong (2013), looking at how financial investors affect risk sharing and information discovery, argue that financialization has substantially changed commodity markets.

1.2 Structure

The paper unfolds into eight chapters. Chapter 2 presents the dataset and how it is obtained, the 3rd one introduces the signals and their derivations. Chapter 4 tries to refining those signals and chapter 5 presents the combination methods of said signals. Chapter 6 is dedicated to the artificial neural network while chapter 7 presents various robustness checks before concluding with chapter 8.

2 Data

2.1 Data Collection

The study is conducted on the daily settlement price and total return of 27 commodity futures contracts over the period of January 2nd, 1990 to December 30th, 2016, downloaded from Datastream. **Table 1** displays the summary statistics of the commodities studies. The sample contains various types of commodities, such as energy (Natural Gas, Brent, WTI, Heating Oil (ULS Diesel), Gasoline); agriculture (Corn, Cotton, Oats, Soybean, Soybean Oil, Soybean Meal, Cocoa, Sugar, Coffee, Rough Rice, Orange, Lumber, Wheat, HRW Wheat); livestock (Live Cattle, Feeder Cattle, Lean Hogs); metals (Silver, Gold, Platinum, Palladium, Copper). For the gasoline futures contract we used the unleaded gasoline daily prices until October 2005, after this date the series switch to the Reformulated Blendstock for Oxygenate Blending (RBOB). Daily closing of the S&P-GSCI and the Bloomberg Commodity Index are instead obtained from Bloomberg. Furthermore, Fama-French 5 Research Factors are downloaded from the Kenneth R. French Data Library website while the inflation data from FRED. Constructing the Hedging Pressure portfolio has been possible using the U.S. Commodity Futures Trading Commission weekly data on the Commitment of Traders, provided in their website.

2.2 Constructing the Strategy

In our strategy, only the most liquid future contracts (nearest to maturity) are traded and held in the portfolio. This is done to reduce to a minimum transaction costs and for liquidity reason. Returns are calculated by holding the closest to maturity future contract one day prior to first notice day and then rolling over to the second-closest contract. First notice day (FND) is the first day that holders of long positions may be informed that they have been assigned a delivery of a future contract. Traders that do not intend to receive (or provide) delivery, as in our case, should close their position before FND. Following the approach adopted in Szakmary et al. (2010), the long-short portfolios are fully collateralized, meaning that all the returns displayed unless explicitly referred as total are excess returns. In other words, the returns of the portfolios are divided by two. Thus, if one was to retain the gross performance, he would simply add the risk-free rate.

3 Individual Signals

3.1 Creating the Portfolios

In each of the portfolios based on the individual signals, we go long on the quintile (20%) that is expected to outperform and go short on the quintile that is expected to underperform. Considering our sample with 27 commodities, we therefore create an equally weighted portfolio of 10 futures (5 long and 5 short), with a 1-month horizon.

3.2 Momentum

Momentum based strategies are one of the most basic signals of technical analysis. In our analysis we examine various methods to calculate the signal, over different rolling windows (RW = 1,3,6,12). Result strongly indicate on the exponentially weighted average based signal adjusted for volatility being the best performing model. **Table 2** sums the performance of the four momentum strategies.

3.2.1 Simple Moving Average

The sorting signal is based on the simple average of past performance over the selected ranking period. Considering the erratic nature of each commodity, there is the much more probability that highly volatile assets (like natural gas) could end up at the far ends of the rankings, while less volatile assets (like gold) are stuck in the middle of the rankings. To address this issue, we adjust by simply dividing for the historical volatility calculated over the rolling window of said commodity.

3.2.2 Exponentially Weighted Moving Average

This trend-following rule follows the same concept as the simple moving average, with the difference that in this case we use an exponentially weighted average, further raising the relative weight of recent observations. We adjust for volatility in this case too.

3.3 Term Structure

When inventories are low, the term structure tends to be downward sloping because the cost of storage is below the convenience yield. Hence, in a backwardated curve, the TS strategy dictates to buy as the price of the future will tend to rise. On the opposite, when inventories are high, the term structure tends to be upward sloping because inventory holders are encouraged to hold the physical commodity and sell forward at a premium. Hence, in a contangoed curve, the TS strategy dictates to sell as the price of the future will tend to decline.

Having in mind the rules for the TS strategy, the corresponding sorting signals is the moving average on each rolling window of the log differential between the nearest and the second-nearest futures contract. The highest average roll-yield commodities are held in the long portfolio and the lowest are held in the short portfolio. **Table 3** presents the results of the TS strategy.

3.4 Hedging Pressure

The hedging pressure hypothesis (or risk transfer) is the oldest hypothesis of a source of a commodity futures risk premium. Keynes(1930) and Hicks(1939) discussed that a risk premium for speculator existed as a reward for accepting the price risk which hedgers sought to transfer.

Hirshleifer(1990) provides an equilibrium-based generalized hedging pressure hypothesis where non-participation effects lead to hedging pressure influencing the risk premium of commodity futures. This theory assumes that risk premiums are present in both backwardated markets (when hedgers are net short and speculators are net long) and in contangoed markets (when hedgers are net long and speculators are net short). Following this idea, we construct the sorting signal for the speculator portfolio and the commercial portfolio as proposed by Basu and Miffre(2013).

The Commitments of Traders dataset provide a breakdown of each Tuesday's open interest on commodity futures contract. The Commodity Futures Trading Commission (CFTC) classifies traders based on the size of their positions into reportable and nonreportable. Reportable traders constitute 70% to 90% of the open interest of any futures markets and are further classified as commercial or non-commercial (speculators). We now define the hedging pressure variable for each category as the number of long open position over the total position in that category.

$$HdgPress_{Spec} = \frac{OI^L_{Spec}}{OI^L_{Spec} + OI^S_{Spec}}$$

$$HdgPress_{Comm} = \frac{OI^L_{Comm}}{OI^L_{Comm} + OI^S_{Comm}}$$

Table 3 presents the results of the Hedging Pressure strategies.

3.4.1 Speculator Portfolio

The speculator-based portfolio is formed by going long on the highest average speculator hedging pressure and going short on the lowest average speculator hedging pressure, on (RW=1,3,6,12). We define this portfolio as *Spec*.

3.4.2 Commercial Portfolio

The commercial-based portfolio is formed by going long on the lowest average commercial hedging pressure and going short on the lowest average commercial hedging pressure, on (RW=1,3,6,12). We define this portfolio as *Comm*.

3.5 Idiosyncratic Volatility

The relation between idiosyncratic risk and returns has been subject of intense study and debate. Sharpe(1964) suggests that idiosyncratic risk should not be priced as investor can diversify away the risk of single assets. Empirical analysis has delivered mixed results. Studies as Fama and McBeth(1973); Bali et al.(2005); Bali and Cakici(2008); Fink et al.(2012); Huang et al.(2010); Han and Lesmond(2011) favors the idea that idiosyncratic risk should not be priced, but others as Malkiel and Xu(2002); Goyal and Santa-Clara(2003); Fu(2009); Garcia et al., 2011) report evidence in favor of a positive correlation between idiosyncratic risk and returns.

The role of idiosyncratic risk in the commodity market was firstly studied by Hirshleifer(1988) in a theoretical model that accounts for trading costs and non-marketability of producers claims. Miffre and Fuertes(2013) showed evidence through an empirical study that idiosyncratic volatility is negatively correlated. **Table 3** presents the results of the IVol strategy.

3.5.1 Benchmark model

Unlike other signals, the idiosyncratic risk signal has to be defined on different benchmarks. Benchmarks will be chosen between traditional benchmarks and fundamental commodity benchmark. The model can be expressed as the following

$$r_{k,d} = \alpha_k + \mathbf{F}_d \beta_k + \varepsilon_{i,d}, \quad d = 1, \cdots D \; days$$

where $r_{k,d}$ is the daily return of the k^{th} commodity future contract, $\mathbf{F}_d$ is the systematic risk premia factors matrix, $\varepsilon_{i,d}$ is the error term, and (α_k, β_k) are the OLS estimated parameters. The regression is iteratively run over the days spanned by a monthly rolling window of (RW = 3,6,12). Following the assumption that idiosyncratic volatility, from now defined as IVol, is negatively correlated with returns, we buy (sell) the assets with the lowest (highest) IVol signal which is obtained as the standard deviation of the residuals of the above mentioned regression.

For the "traditional" factors, inspired by traditional asset pricing model, we consider the equity risk premium ($R_m - R_f$), size and value risk premia (SMB, HML); robust minus weak and conservative minus aggressive (RMW, CMA). We also implement the same concept of equity risk premium to the specificity of the commodity futures market by using as factor the S&P-GSCI and/or the Bloomberg Commodity Index (BCOM).

On the other hand, we use factor that capture the fundamentals of backwardation and contango, precisely term structure portfolio (TS) and hedging pressure portfolio.

We consider all of them and choose the model that gives the best performance via backtesting. The model considering as a sole regressor the S&P-GSCI yields the best return, hence we will use it for the rest of the study.

3.6 Higher Moments

Skewness is a measure of the asymmetry of a probability density function, with negative values indicating that the tail on the left side of the probability density function is longer or fatter than the right side. Conversely, positive skew indicates that the right tail is longer or fatter than the left tail.

Kurtosis is a measure of how "fat" are the tails of the distribution. An higher (lower) value corresponds to fatter (thinner) tails. Having this in mind, one ought to have a higher skewness as possible and a lower kurtosis as possible. **Table 4** presents the results of the TS strategy.

3.6.1 Constructing Skewness Portfolio

We construct the skewness based signal by calculating the skewness of the daily futures contract returns over a monthly rolling window of (RW = 1,3,6,12). We then buy (sell) the contract with the highest (lowest) average skewness signal. The unbiased estimator for the standardized sample skewness is:

$$\frac{\sqrt{n(n-1)}}{n-2}\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^3}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2}\right)^3}$$

3.6.2 Constructing Kurtosis Portfolio

We construct the kurtosis based signal by calculating the kurtosis of the daily futures contract returns over a monthly rolling window of (RW = 1,3,6,12). We then buy (sell) the contract with the lowest (highest) average kurtosis signal. The unbiased estimator for the standardized sample kurtosis is:

$$\frac{n-1}{(n-2)(n-3)}((n+1)k-3(n-1))+3$$

with

$$k=\frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^4}{\left(\sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2}\right)^2}$$

4 Refining Signals

4.1 Refining Signals

In this section we will try to improve our signals. More specifically, we are going to improve the IVol signal and the Momentum signal by forecasting next period conditional volatility and next period total volatility (to adjust our momentum estimate), respectively.

4.2 Forecasting Idiosyncratic Volatility

Miffre & Fuertes(2013) in their empirical study showed negative correlation between returns and idiosyncratic volatility. In their study, they constructed the IVol signal by computing the standard deviation of daily residuals of a given benchmark over a rolling window. We will improve the signal by forecasting the next period non-systematic volatility.

We look at the residuals of each commodity to assess heteroscedasticity by performing the Engle's ARCH test. **Table 5** shows that each of the 27 commodities studied display conditional heteroscedasticity (ARCH effect). We then estimate via maximum likelihood estimation a GARCH(1,1) model.

$$r_{k,d} = \alpha_k + \mathbf{F_d}\beta_k + \varepsilon_{i,d}, \quad d = 1, \cdots D\ days$$

$$\varepsilon_{i,d} = \sigma_{i,d} Z_d$$

Where Z_d is a strong white noise process and $\sigma_{i,d}$ is the conditional volatility of the process, which follows a certain process. **Table 5** shows the summary statistics of the residuals of each commodity. High kurtosis and the Jarque-Bera test confirm the non-normality of the process. Furthermore, negative skewness (a common pattern in financial series) hints to the non-symmetry of the distribution, where negative movement tend to have a bigger impact on the conditional volatility. Having this in mind, we estimate four models over three different distributions.

4.2.1 The Models

GARCH(1,1)

The Autoregressive Conditional Heteroscedasticity (ARCH) model has been proposed in a seminal paper by Engle (1982). The model uses past disturbances in order to model variance of the process. Empirical evidence showed that an high ARCH order has to be chosen to catch the variance of the series. To fix this problem, the Generalized ARCH (GARCH) of Bollerslev (1986) uses previous period variance to reduce the number of estimated parameter from ∞ to 2. The model boils down to:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

with $\alpha_0, \alpha_1, \beta_1 \geq 0$ and $\alpha_1 + \beta_1 < 1$

GJR-GARCH(1,1)

As previously mentioned, financial time series exhibit volatility clustering (large changes are followed by large changes) and fat-tailness, or in more appropriate statistical term, a high value of kurtosis, way above the value of 3 of a Gaussian distribution. Furthermore, Black (1976) noticed the so-called "leverage effect", in which volatility is higher after negative shocks than after positive shocks of same magnitude.

The GARCH model can capture the first two characteristic of financial series but the symmetry of the model cannot capture the third one. To correct this and capture the "leverage effect", various expansions have been proposed, Exponential GARCH of Nelson(1991), GJR-GARCH of Glosten, Jagannathan and Runkle (1993) and the furthest generalization with Asymmetric Power ARCH (APARCH) of Ding, Granger and Engle (1993). The GJR(1,1) boils down to:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 S_{t-1}^- \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

with $\alpha_0, \alpha_1, \beta_1 \geq 0$ and $\alpha_1 + \gamma_1 \geq 1$ and $\alpha_1 + \beta_1 + \frac{1}{2}\gamma_1 < 1$

where S_{t-1}^- is a dummy variable that takes value 1 if the shock is negative or 0 if positive. Note that if $\gamma_1 = 0$ the model is simply a GARCH(1,1).

EGARCH(1,1)

The exponential GARCH (EGARCH) model by Nelson(1991) is another asymmetric GARCH model and has the following representation:

$$\log \sigma_t^2 = \alpha_0 + \gamma_1(|Z_t| - E\,|Z_t|) + \beta_1 \log \sigma_{t-1}^2 + \alpha_1 \log Z_t$$

For Gaussian: $E\,|Z_t| = \sqrt{\frac{2}{\pi}}$

For Student's t: $E\,|Z_t| = \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}$

For Skewed Student's t: $E\,|Z_t| = \frac{2\xi^2}{\xi+\frac{1}{\xi}} \sqrt{\frac{\nu-2}{\pi}} \frac{\Gamma\left(\frac{\nu-1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)}$

The main advantage of the EGARCH model is that since $\log \sigma_t^2$ may be negative, there are less restrictions on the parameters. However, forecast of conditional variances from an EGARCH model are biased. This can be shown through Jensen's inequality.

$$E[\sigma_t^t] \geq \exp\left(E[\log \sigma_t^2]\right)$$

APARCH(1,1)

The Asymmetric Power ARCH model by Ding, Granger and Engle (1993) generalizes the GJR model by introducing an additional parameter which models the exponent. In practical terms, if δ is equal to 1, the APARCH models the volatility and if equals to 2 it models the variance.

$$\sigma_t^\delta = \alpha_0 + \alpha_1(|\varepsilon_{t-1}| - \gamma_1 \varepsilon_{t-1})^\delta + \beta_1 \sigma_{t-1}^\delta$$

4.2.2 The distributions

The GARCH models are estimated through a maximum likelihood approach. The ML approach as the words say is based on maximizing the likelihood function with respect of unknown parameters. The likelihood function is the joint density of ε_t for given parameters of θ, where $\theta = (\alpha_0, \alpha_1, \beta_1)$ for the case of GARCH(1,1). Assuming the independence of the Z_t process, maximizing the joint density is the same of maximizing the product of the marginal densities. Furthermore, by monotonicity of the log function the maximizer of the product is equivalent to the sum of the logs.

As we discussed in the previous section the non-normality of the residuals, it is useful to consider other distributions than the Gaussian, hence we will consider three distributions: Normal, Student-t (in which we will estimate also the degrees of freedom of the parameter, ν) and the Skewed Student-t (in which we will estimate also the degrees of freedom of the parameter, ν, and the asymmetric parameter, ξ).

Normal

The normal distribution is the most used in estimating a GARCH model. The log-likelihood function is given by

$$L = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\log(\sigma_t^2) - \frac{1}{2}\sum_{t=1}^{T}\left(\frac{\varepsilon}{\sigma}\right)^2$$

where T is the number of observations.

Student-t

The student T is the generalization of the Normal Distribution, in which the parameter ν controls the thickness of the tails. As $\nu \to \infty$, the t-distribution approaches the normal distribution with mean 0 and variance 1. The log-likelihood function is given by

$$L = T\Bigg(\log\left[\Gamma\Big(\frac{\nu+1}{2}\Big)\right] - \log\left[\Gamma\Big(\frac{\nu}{2}\Big)\right] - \frac{1}{2}\log\Big[\pi(\nu-2)\Big]\Bigg)$$

$$-\frac{1}{2}\sum_{t=1}^{T}\Bigg[\log(\sigma_t^2)+(1+\nu)\log\Big(1+\frac{\epsilon^2}{\sigma_t^2(\nu-2)}\Big)\Bigg]$$

where again ν is the degrees of freedom, $\nu > 2$ and $\Gamma\Big(.\Big)$ is the gamma function.

Skewed Student-t

Skewness is an important factor in financial application. Therefore a distribution that can model the asymmetry is quite important. Lambert and Laurent (2001) have extend the Skewed Student density to the GARCH framework. The log-likelihood function is given by

$$L = T\Bigg(\log\Bigg[\Gamma\Big(\frac{\nu+1}{2}\Big)\Bigg]-\log\Bigg[\Gamma\Big(\frac{\nu}{2}\Big)\Bigg]-\frac{1}{2}\log\Bigg[\pi(\nu-2)\Bigg]+\log\Bigg[\frac{2}{\xi+\frac{1}{\xi}}\Bigg]+\log\Bigg[s\Bigg]\Bigg)$$

$$-\frac{1}{2}\sum_{t=1}^{T}\Bigg[\log(\sigma_t^2)+(1+\nu)\log\Big(1+\frac{(sz_t+m)^2}{\nu-2}\xi^{-2I_t}\Big)\Bigg]$$

Where ν is the degree of freedom, $\Gamma\Big(.\Big)$ is the gamma function, ξ is the asymmetry parameter, $m = \frac{\Gamma\big(\frac{\nu+1}{2}\big)\sqrt{\upsilon-2}}{\sqrt{\pi}\Gamma\big(\frac{\nu}{2}\big)}\Big(\xi-\frac{1}{\xi}\Big)$, $s = \sqrt{\Big(\xi^2+\frac{1}{\xi^2}-1\Big)-m^2}$ and

$$I_t = \begin{cases} +1, & \text{if } z_t \geq -\frac{m}{s} \\ -1, & \text{otherwise} \end{cases}$$

Notice that when ξ equals 1, the function boils down to the symmetric student-t.

4.3 Refining Momentum

In the previous chapter we discussed that the momentum strategy section can be improved by correcting with the volatility calculated over a rolling window of past returns. The idea in refining this strategy is to forecast next period (on a monthly basis) volatility so to update the correction term with the forecasted volatility. Basically we take the average of the standard deviation of the past 11 and 5 months with the forecasted next period volatility. More specifically, we will implement this method to the best performing momentum strategy, the exponentially weighted momentum corrected by the volatility. Moreover, given the fact that the forecasting model necessitates a starting rolling window over which to estimate the parameters, for the first 60 months (the initial rolling window) the new "refined" strategy mimics the old strategy and then implements the new strategy.

4.3.1 The Forecasting Model

In order to forecast next period (monthly) volatility we implement a model inspired by a combination of a GARCH model and the HAR (Heterogeneous Auto-Regressive) by Corsi (2009). As previously discussed, the idea of the ARCH model is that squared past returns influence the next period volatility. Moreover, the GARCH introduces autocorrelation in the volatility process. The partial autocorrelation function of the equally weighted portfolio monthly volatility in **Figure 2** confirms this claims, suggesting an autoregressive process of order 2. The more recent HAR propose a simple but effective method to model the asymmetric propagation of volatility. The idea is that volatility over longer time intervals have stronger influence on those shorter time intervals, possibly because long term volatility matters for short term traders while short time volatility does not affect long term strategies. Corsi in the simple HAR models this fact by regressing volatility calculated over 3 different horizons through a cascade model. The HAR model is built for high frequency data and uses as an estimate for the volatility the Realized Volatility, which is the sum of squared returns. This two are the differences between the original model and our application. The HAR of 2009 implements high frequency data, while we are using daily data. In fact, HAR computes realized volatility over daily, weekly and monthly period; while our time frame is over monthly, semiannual and annual. Furthermore, as an estimate for volatility we will use the standard deviation of daily returns. We now present the HAR model (in our framework) and the final model that we use for forecasting. We compute the average between the 27 commodities of the Mean Squared Error of our model and the ARMA(2,1). Our model has an average MSE of 0.000068 while the ARMA(2,1) has a average MSE of 0.000104.

$$M = monthly, B = biannual, A = annual$$

$$\sigma_t^A = c^A + \phi^A \sigma_{t-1}^A + \upsilon_t^A$$

$$\sigma_t^B = c^B + \phi^B \sigma_{t-1}^B + \gamma^B E_t[\sigma_t^A] + \upsilon_t^B$$

$$\sigma_t^M = c^M + \phi^M \sigma_{t-1}^M + \gamma^M E_t[\sigma_t^B] + \upsilon_t^M$$

with $\sigma_t^A, \sigma_t^B, \sigma_t^M$ representing the standard deviations of daily returns over a a window of one year, six months and one month, respectively. υ represents the error of each regression. Moreover, by straightforward recursive substitution

$$\sigma_t^M = c + \beta^M \sigma_{t-1}^M + \beta^B \sigma_{t-1}^B + \beta^A \sigma_{t-1}^A + \varepsilon_t$$

Now the combined model is presented

$$\sigma_t^M = \underset{(2.37)}{\beta_0} + \underset{(5.78)}{\beta_1^M} \sigma_{t-1}^M + \underset{(3.13)}{\beta_2^M} \sigma_{t-2}^M + \underset{(2.31)}{\beta^B} \sigma_{t-1}^B + \underset{(4.37)}{\beta^A} \sigma_{t-1}^A + \underset{(1.97)}{\psi} r_{t-1}^2 + \eta_t$$

with r_{t-1}^2 representing past period squared returns and the numbers in the parenthesis under the coefficients are the average t-test taken from the 27 commodities studied.

4.4 The Results

The refined signals display a robust indication that forecasting total monthly variance and conditional volatility through a GARCH model can improve the Mom and

IVol strategy performance, respectively. In fact, the Momentum strategy with forecasted volatily is superior in both the 6 months and 12 months rolling windows. Similar argument can be done for the IVol refined strategy, in which all models using any variation of GARCH and distribution are superior of the standard IVol model. **Table 6** and **Figure 3** display the results.

5 Combining the Signals

5.1 Studying the interaction between signals

In order to justify the combined use of the different signals we must study the interaction between them and prove that they do not contain the same information.

In order to provide evidence, we compute the correlation matrix with the portfolios created using the single signals. The return correlations range between -0.11% to 0.18%. Except for the higher moments based strategies, all these portfolios are deemed to capture the fundamentals of contango and backwardation, hence we expect, when the correlation is statistically significant, positively correlated portfolio returns. This is indeed the case for all cases. Moreover, unlike the hedging pressure hypothesis (Cootner,1960) and theory of storage (Working,1949), Momentum strategy has no theoretical background. Despite this, empirical studies (Gorton et al,2012) showed that recent winners exhibit positive roll-yield, low inventories and net short hedging. **Table 7** displays the correlations between portfolios.

5.2 Construction of the combined strategies

To construct the sorting signal for the combined strategies, we assign a score to the position of every asset in the single signal sorting portfolios, going from 1 (the most outperforming) to N (the most underperforming), following the screening strategy proposed by Achour *et al*(1998). We then add together this "score matrices" and we sort again in ascending order, hence the 20% assets with lowest score will be the in the long portfolio and conversely the 20% assets with the highest score will be in the short portfolio. Before combining all the signals, we combine the Hedging Pressure portfolios (Commercial and Speculative Pressure) and the Higher Moments portfolios (Skewness and Kurtosis) in order to reduce the total number of portfolios to combine to avoid overlapping.

5.2.1 Combined Hedging Pressure

As explained in the previous section, to combine 2 or more strategies we assign a score from 1 to N to each commodity in the each sorting signal. For the Combined Hedging Pressure the highest score of N is given to the lowest commercial hedging pressure and 1 to the highest average commercial hedging pressure. Conversely, the highest score of N is given to the highest speculator hedging pressure and 1 to the lowest average speculator hedging pressure. We sort commodities on their total score and go long on the commodities with the highest score and short on the commodities with the lowest score.

5.2.2 Combined Higher Moments

In a similar fashion to the Combined Hedging Pressure portfolio, we combine Skewness and Kurtosis by assigning the highest scores to the commodity with the highest skewness and lowest kurtosis, and lowest scores to the commodity with the lowest skewness and highest kurtosis. We sort commodities on their total score and go long on the commodities with the highest score and short on the commodities with the lowest score. **Table 4** displays the performance of the two combined strategies.

5.3 Total Combined

Having shown the non-overlapping power of the signals for commodity futures returns, we combine the Mom, TS, IVol, Combined Hedging Pressure and Combined Higher Moments signals. **Figure 4** plots the cumulative (excess) returns of the five long-short strategies to be combined and the total combined long-short strategy. The graph suggests that combining the five signals adds value relative to exploiting each signal individually. In **table 8** one can examine the performance of the total combined strategy. Furthermore, each of the long-short portfolios (single and combined) yield a significant and substantial improvement over the S&P-GSCI. It is paramount to notice that this combination strategy equally weights each component.

5.4 Linear Regression based Combination

We recognize that signals do not contribute in the same way to the prediction of commodities performance, as Mom and TS seems to be the best performing having a higher Sharpe ratio between all of the other strategies. Furthermore, performance of the strategies varies through time. To account for this, we regress the signals for in a simple OLS estimation against the next period returns.

$$r_{t+1}^{k} = \mathbf{S}_t^k \beta' + \varepsilon_{t+1}$$

$$\mathbf{S}_t^k = \begin{bmatrix} \mathbf{1} & Mom_t^k & TS_t^k & IVol_t^k & \left(HdgPress_{Spec,t}^k - HdgPress_{Comm,t}^k\right) & \left(SK_t^k - KU_t^k\right) \end{bmatrix}$$

$$\beta = \begin{bmatrix} \underset{(3.41)}{\beta_0} & \underset{(1.15)}{\beta_1} & \underset{(6.87)}{\beta_2} & \underset{(5.36)}{\beta_3} & \underset{(4.17)}{\beta_4} & \underset{(1.07)}{\beta_5} \end{bmatrix}$$

where r_{t+1}^{k} is the next monthly return of the k commodity; $\mathbf{S}_t^k$ is the matrix containing the signals of the k commodity and β the vector containing the loading factors with the average t-test in the below parenthesis. Notice that when we define the elements of the $\mathbf{S}_t^k$ matrix $(Mom_t^k, TS_t, ...)$, we intend $Mom, TS, ...etc$ as the signals (past averaged performance, the roll-yield), not the returns based on that signal. Furthermore, to account for the combination of the Hedging Pressure signals and the Higher Moments signal, we create a unique factor by taking the difference of the two factors.

5.5 Risk-Return based Combination

Regressing noisy signals with noisy data as the commodities future returns can produce inaccurate result. Furthermore, not accounting for the non-linearity of the process could result in big forecasting errors. We appreciate the simplicity of the combination assigning values to each commodity on each signal of the previous section. In said strategy, each signal is given an equal weight. We try to change this by giving a weight to each signal based on their return-risk performance when implement as single signal portfolio strategies. In other words, the value we assign at each commodity based on their ranking, it is going to be multiplied by a scalar. The scalar for each strategy is chosen in the following way. We compute, using an expanding rolling window with starting length of 12 observations, the Sharpe and Sortino ratio of each single signal portfolio strategy. The weights for the sorting signal for the Risk-Return based Combination are given by

$$w_i = \frac{SR_i}{\sum_i SR_i}$$

with $i = Mom, TS, IVol, CombinedHedgingPressure, CombinedHigherMoments$ and SR_i is either the Sharpe ratio or the Sortino ratio of the i single signal portfolio strategy. It is trivial to see that $\sum_i w_i = 1$.

Among the two risk-return combination strategies, the Sortino ratio seems to perform better. Furthermore, this method yields the best performance over the all combining strategies studied so far. **Table 8** and **Figure 4** displays the result of the combining strategies presented so far.

6 Neural Networks

6.1 Machine Learning

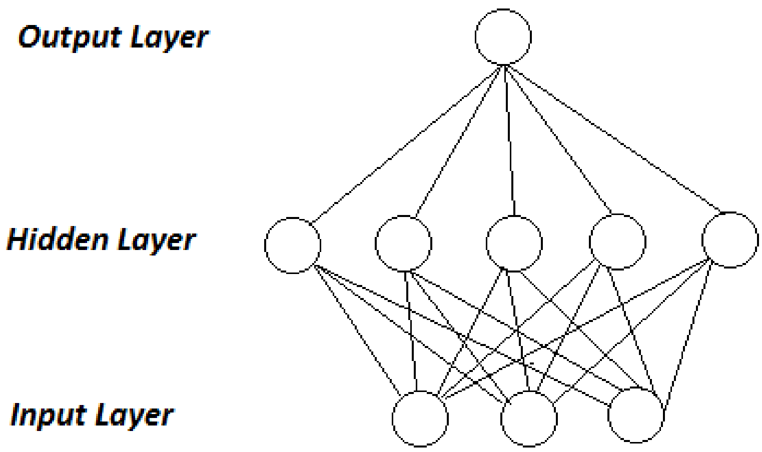
Machine learning represent a group of computational methods strictly related to pattern recognition. Neural networks, adaptive filters are example of machine learning methods which give "computers the ability to learn without being explicitly programmed", as Arthur Samuel in 1959 defined them. Although the application of Neural Network is quite recent, the theory behind it goes back to the famous paper of 1950 "Computing Machinery and Intelligence" by Turing. In his proposal the English mathematician explores the various characteristics that could be possessed by a thinking machine and the various implications in constructing one. In 1951, the first neural net machine (the SNARC) was designed and built by Marvin Minsky. Another important step happened in the 1980s with the rediscovery of backpropagation, which is a training method for artificial neural networks. In the 2000s, thanks to advances in computational speed, deep learning becomes feasible and neural networks see widespread commercial use.

6.2 A Simple Algorithm

In this section we present a simple 3 layers (1 hidden layer) neural network algorithm. The table below presents the problem that we will try to forecast. More specifically, the objective is to forecast the output at $t = 8$ using the data from $t = 1$ to $t = 7$. Notice that in the data the third column is irrelevant and the first and second column behave like a XOR gate, in which the output is true only if one, and only one, of the inputs is true.

t	Inputs			Output
1	0	0	1	0
2	0	0	0	0
3	1	0	1	1
4	0	1	1	1
5	1	1	1	0
6	1	0	0	1
7	0	1	0	1
8	1	1	0	0

The problem is highly non linear. To show this we try to solve the problem by solving a system of linear equations, $Ax = b$, for x. The solution $x = A^{-1}b$ returns the x weights and we multiply them by the input at $t = 8$. The weights (x) are equal to $[0.5 \quad 0.5 \quad 10^{-17}]$. The weights displays that solving the problem in this way can capture the irrelevancy of the third column, but not the non-linearity of the XOR gate. In order to account for this, we build a Neural Network of three layers, with the hidden layer of 5 neurons.



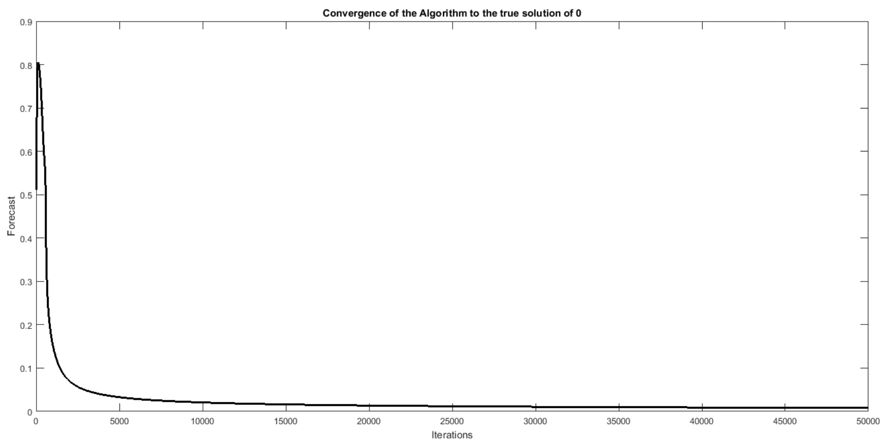
The neuron output (the output that the Neural Network estimates), is the weighted sum of the previous layer. This value is normalized with a so-called activation function. In our specific case (output between 0 and 1), we choose the Sigmoid function.

$$S(x) = \frac{1}{1 + e^{-x}}$$

A nice property of the Sigmoid function is its gradient, which can be easily defined as:

$$S^{'}(x) = \frac{e^{-x}}{(1+e^{-x})^2} = \frac{1}{1+e^{-x}}\frac{e^{-x}}{1+e^{-x}} = \frac{1}{1+e^{-x}}\Big(1 - \frac{1}{1+e^{-x}}\Big) = S(x)\Big(1 - S(x)\Big)$$

We can now start our algorithm. We initialize the weights with random uniform numbers between -1 and 1. We calculate the error between the predicted output and the actual output and work our way backwards. Firstly by computing the hidden layer delta, which is the error multiplied by the sensitivity of a neuron to an input, which is the gradient of the Sigmoid function. We now propagate the error backward, multiplying the hidden layer delta by the weights of the hidden layer. The delta of the input layer is calculated by multiplying the propagated error by the gradient of the Sigmoid function calculated with the output of the input layer. We can now adjust the weights of the two layers, multiply the inputs of the training set (data from $t = 1$ to $t = 7$) by the input layer delta for the input layer and for the hidden layer multiply the output of the input layer by the hidden layer delta. We repeat this process for a large N. The figure below plots the forecasted value over N iterations. We can clearly see that already at 5000 iterations the forecasted value is around 0.05, very close to the true value of zero; ending up at 0.0076 after 50000 iterations.



6.3 Application in Commodity Forecasting

We now return to our task of forecasting commodities returns. In the previous chapter we showed that our signals are able to capture risk premia and generate positive returns. Furthermore, we proved that the combination of these signals can improve the performance of single signals strategies. Moreover, their contribution to returns is not equally weighted. In fact, the regression based combination and the Sortino Ratio based combination both beat the equally-weighted signals combination. We implement the neural network method to address the probable non-linear contribution of the signals to returns.

6.3.1 Normalizing the signals

To stabilize the problem, we transform the signal matrix of section 5.4 into a -1,0,1 input matrix. We achieve this by assigning the value of 1 to the commodities that are "signalled" to be in the long portfolio by each signal, -1 if they are in the short portfolio and 0 if in neither of them. Notice that for the regression based combination we subtracted the Hedging Pressure Speculator signal with the Hedging Pressure Commercial signal to recreate the combined Hedging Pressure signal. A similar argument is made for the Higher Moment signal. For our neural network case, we can directly look at the Combined Hedging Pressure and the Higher Moment signal to construct our 1,0,-1 signal matrix.

6.3.2 Levenberg-Marquardt Algorithm

For more complicated problems we rely on a more sophisticated optimization algorithm. Like the quasi-Newton methods, the Levenberg-Marquardt algorithm is designed to approach second-order training speed without having to compute the Hessian matrix. In fact, when the performance function has the form of a sum of squares (as is typical in training feedforward networks), then the Hessian matrix can be approximated as:

$$\mathbf{H} = \mathbf{J}^T\mathbf{J}$$

The minimization problem reads as follows:

$$\underset{\beta}{argmin} \sum \Big(y_i - f(x_i, \beta)\Big)^2$$

To initialize the minimization process, initial guess has to be inputted. In each iteration step, β is replaced by a new estimate $\beta + \delta$ which when plugged in the function can be approximated with a first-order approximation:

$$f(x_i, \beta + \delta) = f(x_i, \beta) + J_i \delta$$

Where J_i is the Jacobian. Using matrix notation, we show the optimal value of δ.

$$\left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta}) - \mathbf{J}\boldsymbol{\delta}\right]^T \left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta}) - \mathbf{J}\boldsymbol{\delta}\right] =$$

$$= \left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta})\right]^T \left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta})\right] - \left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta})\right]^T \mathbf{J}\boldsymbol{\delta} - \left[\mathbf{J}\boldsymbol{\delta}\right]^T \left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta})\right] + \boldsymbol{\delta}^T \mathbf{J}^T \mathbf{J}\boldsymbol{\delta}$$

Taking the derivative wrt δ yields

$$\delta = \left[\mathbf{J}^T \mathbf{J}\right]^{-1} \mathbf{J}^T \left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta})\right]$$

Levenberg's contribution is to replace this equation by a "damped version"

$$\delta = \left[\mathbf{J}^T \mathbf{J} + \mu \mathbf{I}\right]^{-1} \mathbf{J}^T \left[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta})\right]$$

where **I** is the identity matrix and μ the dampening factor. When the scalar μ is zero, this is just Newton's method, using the approximate Hessian matrix. When μ is large, this becomes gradient descent with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, μ is decreased after each successful step (reduction in performance function) and is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm. The algorithm stops if the maximum number of iteration is reached, or if the MSE is below a certain threshold value, or the gradient falls below a certain threshold and finally if μ exceeds a certain threshold value. In our neural network problem, the weights will be adjusted with the formula above and the activation function for the output layer (the function that maps the hidden layer to the output layer) will be the hyperbolic tangent, as it ranges from -1 to 1, which describe the range of returns.

6.3.3 The Network

We construct a 3 layers neural network, with the input layer consisting of the 1,0,-1 signal matrix, while the hidden layer consists of a 10 neuron layer. Finally the output layer corresponds to single return vector. We use 250 observations (months) to train our model. We then expand the rolling window as new information becomes available. This means that we have 62 observations out-of-sample in which we will test our model. The NN (neural network) portfolio is constructed using the 250 first returns of the sortino ratio based combinined portfolio and then using the estimates originated from the neural network model to sort the best to worst commodities and, as usual, going long with the predicted best 5 future contracts and short the predicted worst 5 futures contract.

6.3.4 The Results

Figure 5 and **Table 8** show that the neural network provides a better performance during the last 62 months sample. Another indicator of the goodness of the neural network is the correlation between the predicted returns and the actual returns in the 62 months sample against the linear regression based forecasting. In practical terms, this is a measure which indicates how often our model gets right the sign of the return. The neural network has an average of 43.21% average correlation coefficient between commodities while the linear regression model has only a 12.73% average correlation coefficient.

7 Diversification and Robustness Analysis

7.1 Risk Diversification

One of the reasons of the financialization of commodities over the last decade is the risk diversification that they provide to investors. We now address this issue by looking on how our best performing portfolio (the neural network combination) correlates with the S&P500. **Figure 6** plots the rolling correlation (with a window of 60 months) and the unconditional (whole sample) correlation. The correlation for the whole sample is 0.0843 and it is statistically not significant, while the S&P-GSCI has a correlation of 0.15 and it is statistically significant. The rolling window correlation coefficient ranges from 0.3 to -0.2, but with a low number of values around this boundaries, confirming the risk diversification properties of the long-short portfolio. We can conclude that equity investors seeking to diversify their investment should opt for a long-short approach in futures commodity markets.

7.2 Inflation Hedging

The first step in this analysis is to compute the shocks to inflation. The model fitted to the data is a ARMA(1,2) process, as we have chosen the appropriate lags by choosing the smallest value of the Bayesian Information Criterion over 16 different models. Then the residuals of the chosen model are considered as inflation shocks. Correlation between said inflation shocks and our portfolio returns is 0.0071, thus providing a good hedge against inflation. In contrary, the S&P-GSCI shows a correlation coefficient of 0.24 it is statistically significant.

7.3 Robustness Analysis

We now present additional tests to confirm the superiority in performance of the long-short portfolio.

7.3.1 Is the Financialization of Commodities the Reason behind Performance?

In order to assess if the the flow of cash in the commodity futures markets explains the profits of our strategy we present two tests.

Random Portfolio

We create a random long-short portfolio for every month of the sample from January 2000 to June 2008 sorting the commodities using draws from a simulated uniform distribution. We then compute the Sharpe Ratio of this random portfolio. We simulate the above process using a simple Monte Carlo simulation a 1000 times, storing the results of the Sharpe Ratio at every iteration. We should then approach a Normal Distribution with mean zero and variance σ^2 if the financialization has no effect on a random portfolio. The average value of the sharpe ratio is 0.0033, which ticks off the first requirement of a mean equal to zero. To prove normality we perform the Jarque-Bera test, which returns a value of 4.1239, which is below the critical value of 5.9282, and confirms that the distribution as an expected skewness and kurtosis as the normal distribution. Performing the one-sample Kolmogorov-Smirnov test confirm the result of the JB-test, thus we can infer that our distribution of simulated Sharpe Ratio is a normal with zero mean.

Granger Causality

We use the notion of causality proposed by Granger(1969) to test if the increase in open interest position granger causes the returns of our long-short portfolio. We test this with the following regression using monthly data:

$$r_t^p = \underset{(2.73)}{\delta_0} + \underset{(1.79)}{\delta_1} \log(\sum_i OI_{t-1}^i) + \underset{(0.87)}{\phi} r_{t-1}^p + \eta_t$$

where r_t^p is the long-short neural network portfolio and $\sum_i OI_{t-1}^i)$ sum of the open interest of each commodities. The t-test in parenthesis below δ_1 coefficient show us that the movement of the open interest has no impact on performance.

7.3.2 Transaction Cost

Locke & Venkatesh (1997) estimated that futures trading costs range between 0.0004% and 0.033%. As the commodities futures contract studies in this paper are often liquid, cheap and easy to short sell, transaction cost are unlikely to affect performance. Furthermore, the strategy trades with the closest-to-maturity, often the most liquid one. Thus, without doing formal tests, we can infer with reasonable confidence, that transaction cost do not wipe out the abnormal performance of our portfolio.

7.3.3 Liquidity Risk

In the previous subsection we discussed the highly liquidity feature of our long-short portfolio. Nevertheless, we want to test if the outperformance of the combination strategy is merely a compensation for holding illiquid futures contracts. We tackle this issue by calculating the α of the portfolio via a two-factor model formed by the S&P-GSCI and a liquidity risk premium portfolio (LRP) constructed following the idea Pastor & Stambaugh's paper of 2003, and the commodity implementation of Fuertes & Miffre (2014). The regression is the following

$$r_t^p = \underset{(9.87)}{\alpha} + \underset{(3.21)}{\beta}\, S\&P - GSCI_t + \underset{(1.81)}{\gamma}\, LRP_t + \nu_t$$

The estimated significant α is 8.39%, indicating that our long-short neural network strategy is not a compensation for the illiquidity of certain futures contracts.

8 Conclusion

This paper relies on signals that have been shown to predict commodity futures returns in the recent literature. As previous literature indicated, the best performing benchmark for the idiosyncratic volatility is the S&P-GSCI. We go further by enhancing the momentum and idiosyncratic volatility signals forecasting the next period volatility. We further expand upon current literature by presenting different methods for combining the signals, concluding with a combination method based on a artificial neural network. We showed that over the period from 1990 to the end of 2016 an investor that buys commodities with recent high past performance, high average roll-yield, low idiosyncratic volatility, low commercial pressure and high speculator pressure, highest past skewness and lowest kurtosis; and sell commodities that present the opposite above features can obtain an average (between the 5 combination strategies) Sharpe ratio of 1.23. This value is much higher then the Sharpe ratios obtained using the single signals, even more for the S&P-GSCI that stands at 0.11. Moreover, we also proved the diversification and inflation hedging properties of our strategies.

The main goal of this paper was to study the effectiveness of combining the studied signals and the strategies on how to combine them. An open question remains whether adding as a signal inventory levels could improve performance.

A Tables

The table presents the summary statistics of the monthly returns of the 27 commodities studied

	Mean	Volatility	Skewness	Kurtosis	Minimum	Maximum
Natural Gas	0.0074	0.1533	0.1604	4.2431	-0.4937	0.6207
Brent	0.0008	0.0870	0.1487	5.0575	-0.3011	0.3856
WTI	0.0020	0.0900	0.1324	4.3405	-0.3144	0.3683
Heating Oil	0.0070	0.1085	0.8476	6.8686	-0.3382	0.5494
Gasoline	0.0106	0.0938	-0.4221	5.0436	-0.4369	0.3225
Corn	-0.0013	0.0830	-0.6461	5.3737	-0.4363	0.2310
Cotton	0.0008	0.0909	-0.6852	6.0420	-0.4920	0.2688
Oats	-0.0072	0.0974	-0.1602	3.6369	-0.3477	0.3255
Soybean	0.0005	0.0747	-1.0472	7.4846	-0.4403	0.2111
Soybean Oil	-0.0017	0.0739	-0.5726	5.5939	-0.3259	0.2211
Soybean Meal	0.0022	0.0862	-0.6951	6.3281	-0.4201	0.2851
Cocoa	0.0068	0.0874	0.0516	4.2422	-0.3616	0.3005
Sugar	0.0023	0.1004	0.3021	4.9100	-0.3712	0.4686
Coffee	-0.0015	0.1022	0.4486	4.3536	-0.2986	0.4142
Rough Rice	-0.0011	0.0857	-0.1213	5.4157	-0.3899	0.3858
Orange	-0.0040	0.0955	0.3121	3.4323	-0.2646	0.3406
Lumber	0.0008	0.0983	-0.0585	3.3487	-0.2852	0.3265
Wheat	-0.0012	0.0837	-0.0804	3.4330	-0.2692	0.2655
Wheat HRW	-0.0002	0.0784	0.0096	3.8645	-0.2785	0.2595
Live Cattle	0.0054	0.0507	-0.0203	4.0169	-0.2039	0.1810
Feeder Cattle	-0.0013	0.0447	-0.9468	7.0679	-0.2708	0.1206
Lean Hogs	0.0007	0.1030	-0.1643	4.4450	-0.4580	0.3384
Silver	0.0004	0.0831	-0.3340	4.5496	-0.3460	0.3056
Gold	0.0032	0.0464	-0.2087	5.3045	-0.2178	0.1702
Platinum	-0.0005	0.0678	-2.5767	31.4942	-0.6627	0.3189
Palladium	0.0023	0.1012	-0.1114	7.9458	-0.4766	0.5229
Copper	0.0025	0.0697	-0.6760	7.4356	-0.4131	0.2234

FIGURE A.1

The table presents summary statistics for the fully collaterized excess returns for the momentum strategies.
The first strategies is a simple equally weighted average, the second one is the same as the first but correcting for the volatility. The third is the exponential weighted average, while the last one is as the third one but correcting for volatility.

	Momentum simple avg				**Momentum simple avg vol corrected**				**Momentum exp weighted avg**				**Momentum exp weighted avg corrected vol**			
	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12
Long Port Mean	-0.0079	0.0103	0.0043	0.0063	0.0124	0.0033	-0.0042	0.0071	N/A	N/A	0.0117	0.0147	N/A	N/A	0.0146	0.0235
Short Port Mean	0.0117	0.0109	0.0159	0.0053	0.0094	0.0155	0.0180	0.0037	N/A	N/A	0.0013	0.0013	N/A	N/A	-0.0022	-0.0189
Long-Short Port Mean	-0.0196	-0.0006	-0.0116	0.0010	0.0030	-0.0122	-0.0222	0.0034	N/A	N/A	0.0104	0.0135	N/A	N/A	0.0168	0.0423
Long-Short Port GeoMean	-0.0214	-0.0026	-0.0134	-0.0009	0.0012	-0.0138	-0.0234	0.0020	N/A	N/A	0.0010	0.0119	N/A	N/A	0.0159	0.0418
Volatility	0.0635	0.0631	0.0619	0.0597	0.0604	0.0573	0.0527	0.0547	N/A	N/A	0.0636	0.0578	N/A	N/A	0.0528	0.0508
Downside Volatility (0%)	0.0408	0.0376	0.0371	0.0347	0.0326	0.0347	0.0325	0.0335	N/A	N/A	0.0386	0.0325	N/A	N/A	0.0315	0.0268
Skewness	-0.0868	-0.0077	0.1226	0.0677	0.5026	0.0759	0.1988	-0.2996	N/A	N/A	0.1806	0.0338	N/A	N/A	0.4011	-0.1362
Kurtosis	3.8878	3.3640	3.3734	3.2548	4.1747	4.0021	3.6287	3.2883	N/A	N/A	3.6266	3.5814	N/A	N/A	4.1067	3.7407
Maximum	0.0403	0.0424	0.0467	0.0419	0.0466	0.0459	0.0490	0.0393	N/A	N/A	0.0474	0.0420	N/A	N/A	0.0397	0.0418
Minimum	-0.0429	-0.0468	-0.0452	-0.0462	-0.0448	-0.0438	-0.0395	-0.0432	N/A	N/A	-0.0490	-0.0468	N/A	N/A	-0.0456	-0.0417
Sharpe Ratio	-0.0308	-0.0095	-0.1873	**0.0173**	0.0492	-0.2133	-0.4217	**0.0630**	N/A	N/A	0.1642	**0.2328**	N/A	N/A	0.3187	**0.8331**
Sortino Ratio (0%)	-0.4804	-0.0160	-0.3125	0.0297	0.0911	-0.3523	-0.6850	0.1028	N/A	N/A	0.2706	0.4147	N/A	N/A	0.5339	1.5775
Omega Ratio (0%)	-0.5380	-0.0179	-0.3500	0.0332	0.1021	-0.3945	-0.7672	0.1151	N/A	N/A	0.4058	0.6220	N/A	N/A	1.0340	2.0041

FIGURE A.2

The table presents summary statistics for the fully collaterized excess returns of Term Structure, Idiosyncratic Volatility, Speculative Pressure, Commercial Pressure Strategies. All strategies are calculated for rolling windows of 1,3,6,12 months

	Term Structure				Idiosyncratic Volatility				Speculative Pressure				Commercial Pressure			
	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12
Long Port - Mean	0.0215	0.0315	0.0314	0.0254	0.0083	0.0020	0.0048	0.0088	0.0089	0.0050	0.0065	0.0090	0.0054	-0.0019	0.0019	-0.0016
Short Port - Mean	-0.0131	-0.0110	-0.0137	-0.0099	0.0034	-0.0022	-0.0016	-0.0012	-0.0035	-0.0057	0.0034	0.0033	0.0025	-0.0013	0.0036	0.0000
Long-Short Port Mean	0.0345	0.0425	0.0451	0.0352	0.0049	0.0042	0.0064	0.0100	0.0124	0.0106	0.0031	0.0057	0.0030	-0.0006	-0.0017	-0.0016
Long-Short Port GeoMean	0.0334	0.0414	0.0444	0.0341	0.0047	0.0026	0.0049	0.0063	0.0111	0.0092	0.0030	0.0045	0.0017	-0.0019	-0.0030	-0.0029
Volatility	0.0527	0.0542	0.0569	0.0580	0.0548	0.0564	0.0538	0.0520	0.0517	0.0548	0.0548	0.0502	0.0502	0.0510	0.0514	0.0503
Downside Volatility (0%)	0.0238	0.0268	0.0258	0.0268	0.0294	0.0311	0.0303	0.0301	0.0287	0.0328	0.0312	0.0285	0.0306	0.0309	0.0307	0.0293
Skewness	0.6602	0.6002	0.6502	0.5443	0.2286	0.2138	0.0806	-0.1278	0.0824	-0.2869	0.1331	0.0313	-0.2037	-0.2281	-0.0555	0.0452
Kurtosis	4.6730	4.0910	4.9878	3.6603	2.8361	2.9719	3.0489	2.7083	3.4552	3.6651	4.2109	3.4034	4.5774	4.4445	3.7141	3.6701
Maximum	0.0483	0.0449	0.0407	0.0416	0.0450	0.0400	0.0465	0.0446	0.0437	0.0397	0.0435	0.0419	0.0443	0.0415	0.0441	0.0485
Minumum	-0.0485	-0.0403	-0.0414	-0.0440	-0.0401	-0.0437	-0.0445	-0.0431	-0.0456	-0.0391	-0.0403	-0.0442	-0.0465	-0.0488	-0.0484	-0.0395
Sharpe Ratio	0.6556	0.7834	**0.7929**	0.6075	0.0891	0.0746	0.1186	**0.1920**	**0.2397**	0.1943	0.0557	0.1143	**0.0590**	-0.0119	-0.0333	-0.0324
Sortino Ratio (0%)	1.4532	1.5858	1.7481	1.3160	0.1661	0.1353	0.2108	0.3313	0.4316	0.3244	0.0977	0.2014	0.0968	-0.0196	-0.0556	-0.0556
Omega Ratio (0%)	1.5259	1.8237	2.0103	1.5134 #	0.1910	0.1556	0.2424	0.3809	0.4963	0.3730	0.1124	0.2316	0.1114	-0.0226	-0.0640	-0.0639

FIGURE A.3

The table presents summary statistics for the fully collaterized excess returns for the Skewness, Kurtosis, Combined Hedging Pressure, Combined Higher Moments strategies. All strategies are calculated for rolling windows of 1,3,6,12 months

	Skewness				Kurtosis				Combined Hedging Pressure				Combined Higher Moment			
	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12	R = 1	R = 3	R = 6	R = 12
Long Port - Mean	0.0145	0.0174	0.0135	0.0195	0.0020	0.0022	0.0108	0.0027	0.0138	0.0145	0.0050	0.0143	0.0067	0.0081	0.0106	0.0294
Short Port - Mean	0.0040	0.0049	-0.0013	-0.0044	0.0039	0.0004	0.0075	-0.0057	-0.0004	-0.0025	0.0073	0.0087	-0.0048	-0.0047	0.0080	-0.0008
Long-Short Port Mean	0.0106	0.0125	0.0148	0.0239	-0.0019	0.0018	0.0032	0.0085	0.0142	0.0170	-0.0023	0.0056	0.0115	0.0128	0.0026	0.0302
Long-Short Port GeoMean	0.0100	0.0098	0.0136	0.0229	-0.0020	0.0005	0.0032	0.0055	0.0130	0.0157	-0.0037	0.0043	0.0102	0.0111	0.0014	0.0292
Volatility	0.0720	0.0741	0.0508	0.0501	0.0583	0.0515	0.0510	0.0524	0.0508	0.0525	0.0532	0.0513	0.0577	0.0600	0.0541	0.0528
Downside Volatility (0%)	0.0431	0.0414	0.0302	0.0270	0.0302	0.0307	0.0287	0.0293	0.0289	0.0295	0.0316	0.0297	0.0369	0.0337	0.0315	0.0279
Skewness	0.2013	0.2356	-0.2813	0.1796	-0.0480	-0.0857	0.1078	0.0292	-0.0983	-0.0766	-0.0077	-0.0146	0.1030	0.1268	0.0063	-0.0773
Kurtosis	4.2014	4.1312	3.6866	5.0438	3.3210	3.3040	3.1721	2.8418	3.3284	2.9229	3.9869	3.8327	3.8759	3.6339	4.0968	3.1239
Maximum	0.0210	0.0450	0.0513	0.0616	0.0424	0.0456	0.0505	0.0402	0.0453	0.0403	0.0368	0.0403	0.0365	0.0512	0.0415	0.0394
Minimum	-0.0561	-0.0517	-0.0460	-0.0515	-0.0509	-0.0587	-0.0602	-0.0411	-0.0470	-0.0479	-0.0632	-0.0410	-0.0532	-0.0530	-0.0610	-0.0522
Sharpe Ratio	0.1466	0.1686	0.2924	**0.4766**	-0.0324	0.0358	0.0633	**0.1615**	0.2804	**0.3231**	-0.0433	0.1093	0.1993	0.2142	0.0485	**0.5718**
Sortino Ratio (0%)	0.2448	0.3015	0.4920	0.8840	-0.0626	0.0603	0.1127	0.2891	0.4926	0.5748	-0.0729	0.1886	0.3116	0.3807	0.0833	1.0833
Omega Ratio (0%)	0.3673	0.4523	0.7381	1.4617	-0.1002	0.0964	0.1803	0.4625	0.7389	0.8622	-0.1093	0.2828	0.4456	0.5444	0.1191	1.5491

FIGURE A.4

The table displays the higher moments of the residuals computed for the estimation of the IVol signal. Furthermore, the *p*- values of the Jarque-Bera test and the Engle test for heteroskedasticity are reported.

	Skewness	**Kurtosis**	**Jbtest**	**Engle test**
Natural Gas	0.258	11.392	0.001	0.000
Brent	-0.897	17.013	0.001	0.000
WTI	0.604	28.984	0.001	0.000
Heating Oil	-2.283	35.878	0.001	0.000
Gasoline	0.039	9.976	0.001	0.000
Corn	-1.295	27.832	0.001	0.056
Cotton	-0.890	21.218	0.001	0.000
Oats	-1.095	15.580	0.001	0.000
Soybean	-1.014	21.442	0.001	0.000
Soybean Oil	0.130	4.818	0.001	0.000
Soybean Meal	-1.255	15.968	0.001	0.000
Cocoa	0.142	5.617	0.001	0.000
Sugar	-0.328	11.334	0.001	0.000
Coffee	0.168	10.558	0.001	0.000
Rough Rice	0.140	27.253	0.001	0.005
Orange	0.386	11.948	0.001	0.000
Lumber	0.616	11.007	0.001	0.000
Wheat	-0.543	19.397	0.001	0.000
Wheat HRW	-0.236	6.493	0.001	0.000
Live Cattle	-0.901	11.551	0.001	0.000
Feeder Cattle	-0.314	15.316	0.001	0.000
Lean Hogs	0.038	38.148	0.001	0.059
Silver	-0.635	9.781	0.001	0.000
Gold	-0.051	9.938	0.001	0.000
Platinum	-8.428	362.477	0.001	0.000
Palladium	0.801	61.748	0.001	0.042
Copper	-0.217	7.546	0.001	0.000

FIGURE A.5

The table presents summary statistics for the fully collaterized excess returns for the "refined signals".
The first is the momentum strategy with exponentially weighted average and corrected with the forecasted total volatility.
Then 2nd,3rd and 4th strategies are the forecasted idiosyncratic volatilities strategies using a Normal, Student-T distribution and Skewed Student-T distribution.
For each distribution a GARCH(1,1), GARCH-GJR(1,1), EGARCH(1,1) and APARCH(1,1) are fitted to the residuals of the benchmark model with a rolling window of 12 months.

	Momentum exp weighted corrected with forecasted vol				**Normal Distribution**				**Student T distribution**				**Skewed T Distribution**			
	R = 1	R = 3	R = 6	R = 12	GARCH	GJR	EGARCH	APARCH	GARCH	GJR	EGARCH	APARCH	GARCH	GJR	EGARCH	APARCH
Long Port - Mean	N/A	N/A	0.0201	0.0234	0.0093	0.0084	0.0091	0.0093	0.0120	0.0125	0.0111	0.0115	0.0132	0.0133	0.0125	0.0183
Short Port - Mean	N/A	N/A	-0.0135	-0.0255	-0.0051	-0.0071	-0.0005	-0.0091	-0.0038	-0.0068	-0.0010	-0.0068	-0.0019	-0.0066	-0.0068	-0.0007
Long-Short Port Mean	N/A	N/A	0.0336	0.0489	0.0145	0.0155	0.0096	0.0185	0.0159	0.0194	0.0120	0.0184	0.0150	0.0199	0.0194	0.0190
Long-Short Port GeoMean	N/A	N/A	0.0329	0.0487	0.0142	0.0150	0.0088	0.0172	0.0102	0.0112	0.0114	0.0122	0.0099	0.0187	0.0112	0.0178
Volatility	N/A	N/A	0.0506	0.0502	0.0527	0.0527	0.0518	0.0537	0.0538	0.0530	0.0528	0.0526	0.0534	0.0529	0.0538	0.0518
Downside Volatility (0%)	N/A	N/A	0.0260	0.0258	0.0273	0.0273	0.0268	0.0283	0.0295	0.0285	0.0268	0.0290	0.0302	0.0285	0.0285	0.0284
Skewness	N/A	N/A	-0.1034	-0.1516	0.3466	0.3065	0.2190	0.3017	0.0243	0.0263	0.3071	0.0273	0.0265	0.0233	0.0243	0.0557
Kurtosis	N/A	N/A	3.6890	3.4995	3.4293	3.8920	3.7850	3.4830	2.6960	2.8160	3.5320	2.7960	3.3481	2.6945	2.6960	3.3671
Maximum	N/A	N/A	0.0390	0.0428	0.0431	0.0442	0.0423	0.0419	0.0468	0.0443	0.0452	0.0433	0.0451	0.0385	0.0431	0.0452
Minimum	N/A	N/A	-0.0455	-0.0451	-0.0423	-0.0386	-0.0429	-0.0464	-0.0477	-0.0419	-0.0416	-0.0457	-0.0461	-0.0422	-0.0453	-0.0425
Sharpe Ratio	N/A	N/A	0.6645	**0.9733**	0.2745	0.2948	0.1863	0.3439	0.2947	0.3652	0.2279	0.3486	0.2814	**0.3765**	0.3598	0.3671
Sortino Ratio (0%)	N/A	N/A	1.2930	1.8974	0.5292	0.5683	0.3606	0.6517	0.5382	0.6801	0.4497	0.6338	0.4972	0.6988	0.6801	0.6691
Omega Ratio (0%)	N/A	N/A	1.7540	2.0450	0.9261	0.9946	0.6311	1.1404	0.9418	1.1902	0.7870	1.1092	0.8701	1.2230	1.1902	1.1709

FIGURE A.6

The table displays the correlation between each signal. Underlined the p-values are displayed.

	MOM	**TS**	**IVol**	**HDG**	**HM**
MOM		-0.03 *0.64*	0.06 *0.27*	-0.09 *0.10*	0.09 *0.12*
TS			-0.03 *0.66*	0.18 *0.00*	0.02 *0.70*
IVol				-0.04 *0.47*	0.14 *0.01*
HDG					-0.11 *0.05*
HM					

FIGURE A.7

The table presents summary statistics for the fully collaterized excess returns for the combined signals. Moreover, in the last column, the summary statistics for the S&P-GSCI and Bloomberg Commodity (BCOM) indexes are presented. For the benchmark statistics we present the gross return, i.e. the position is not fully collaterized as our strategies.

	Equally Weighed	Regression Based	Return-Risk Weighting		Neural Network	Benchmarks	
	Combined	Combined	Sharpe Ratio	Sortino Ratio	Combined	S&P-GSCI	BCOM
Long Port - Mean	0.0238	0.0353	0.0330	0.0339	0.0439	N/A	N/A
Short Port - Mean	-0.0350	-0.0291	-0.0301	-0.0311	-0.0451	N/A	N/A
Long-Short Port Mean	0.0588	0.0644	0.0632	0.0651	0.0890	0.0237	0.0201
Long-Short Port GeoMean	0.0590	0.0574	0.0605	0.0650	0.0842	0.0036	0.0096
Volatility	0.0523	0.0566	0.0544	0.0546	0.0540	0.2167	0.1453
Downside Volatility (0%)	0.0240	0.0262	0.0262	0.0252	0.0237	0.1367	0.0917
Skewness	0.1270	0.3401	0.0324	0.0624	-0.0845	-0.5211	-0.3728
Kurtosis	3.2369	3.7626	3.1558	3.2684	3.4067	10.9497	7.84
Maximum	0.0412	0.0727	0.0495	0.0496	0.0496	0.0167	0.0124
Mininum	-0.0416	-0.0468	-0.0416	-0.0416	-0.0416	-0.046	-0.0201
Sharpe Ratio	1.1249	1.1373	1.1604	1.1909	1.6487	0.1093	0.13833448
Sortino Ratio (0%)	2.4540	2.4552	2.4137	2.5851	3.7558	0.173372348	0.219193021
Omega Ratio (0%)	2.246	2.0609	2.356	2.4156	3.3379	1.0196	1.034

FIGURE A.8

The table presents the frequency of the futures contract entering the long-short neural network portfolio. Alongside are presented the average OLS estimate of the regression of each commodity against the S&P-GSCI and its t-statistics used to compute the IVol signal

	Long (%)	**Short (%)**	$\beta_{S\&P-GSCI}$	***t*-stat**	Adj R^2
Natural Gas	54.84	6.45	1.1028	6.8175	17.11%
Brent	8.06	6.45	1.4273	30.3055	73.03%
WTI	22.58	41.94	1.5458	31.5345	74.98%
Heating Oil	8.06	16.13	1.4402	24.8298	66.94%
Gasoline	25.81	20.97	0.8847	12.5363	37.64%
Corn	0.00	25.81	0.3226	3.9399	6.55%
Cotton	4.84	14.52	0.1944	2.4094	3.06%
Oats	0.00	4.84	0.3149	2.9531	4.06%
Soybean	16.13	14.52	0.2981	4.1997	7.64%
Soybean Oil	1.61	8.06	0.3008	5.0828	11.82%
Soybean Meal	17.74	11.29	0.2675	3.1741	4.24%
Cocoa	8.06	0.00	0.1640	1.9398	2.39%
Sugar	4.84	1.61	0.2557	2.5553	3.23%
Coffee	25.81	11.29	0.2312	2.4359	3.82%
Rough Rice	37.10	25.81	0.1047	1.4916	1.74%
Orange	17.74	29.03	0.0982	0.9960	0.20%
Lumber	41.94	27.42	0.0736	0.7186	0.24%
Wheat	24.19	19.35	0.3455	3.7162	5.99%
Wheat HRW	29.03	33.87	0.3121	3.7062	6.02%
Live Cattle	19.35	20.97	0.0857	1.6140	1.55%
Feeder Cattle	1.61	3.23	0.0251	0.6258	0.34%
Lean Hogs	22.58	12.90	0.1494	1.1968	0.28%
Silver	37.10	30.65	0.4199	4.5421	9.22%
Gold	8.06	12.90	0.2055	4.2352	8.25%
Platinum	19.35	45.16	0.2027	3.1502	5.43%
Palladium	30.65	37.10	0.2471	2.8532	5.69%
Copper	12.90	17.74	0.3504	4.9865	11.98%

FIGURE A.9

B Figures

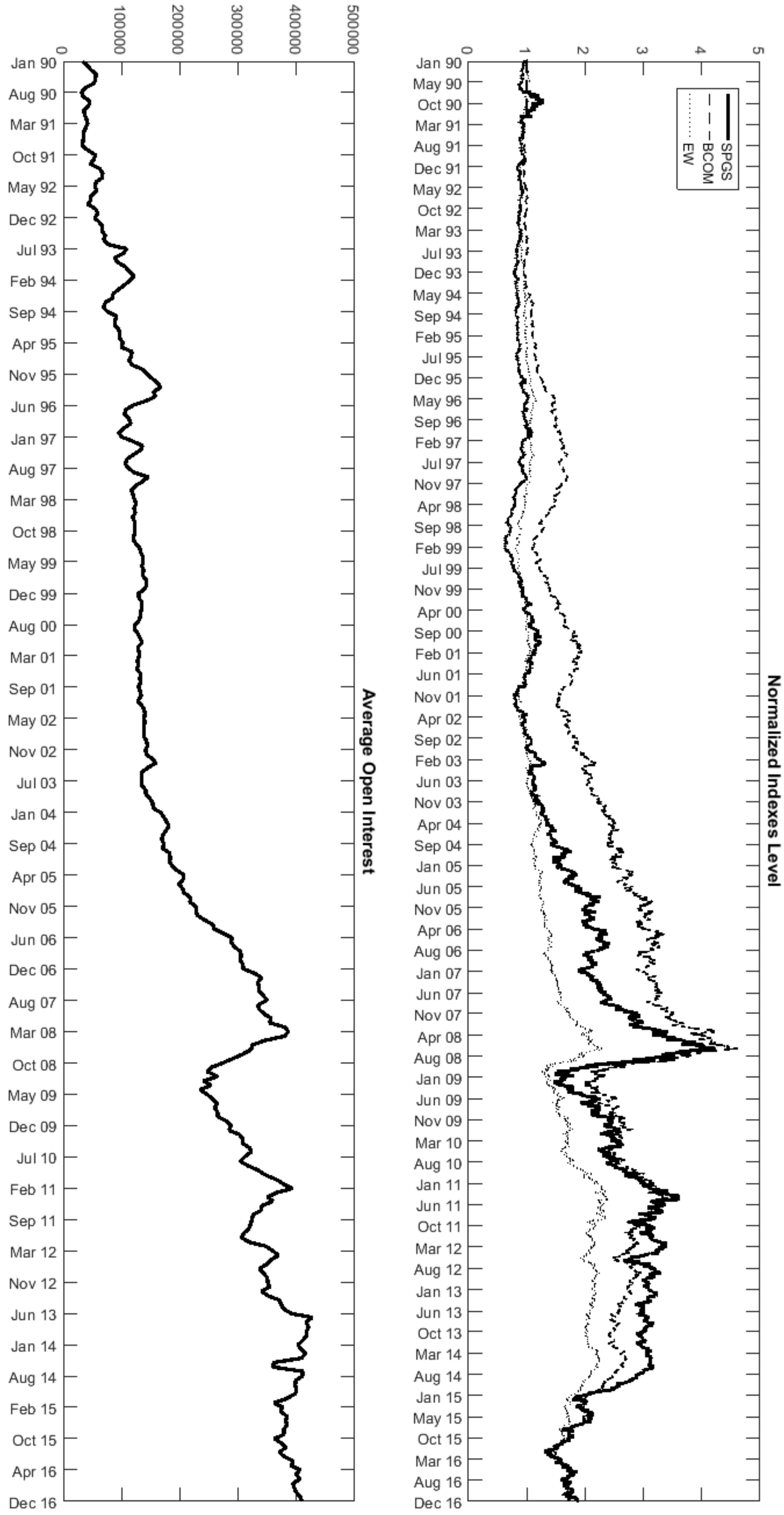

Normalized Indexes Level
SPGS
BCOM
EW
Average Open Interest

FIGURE B.1

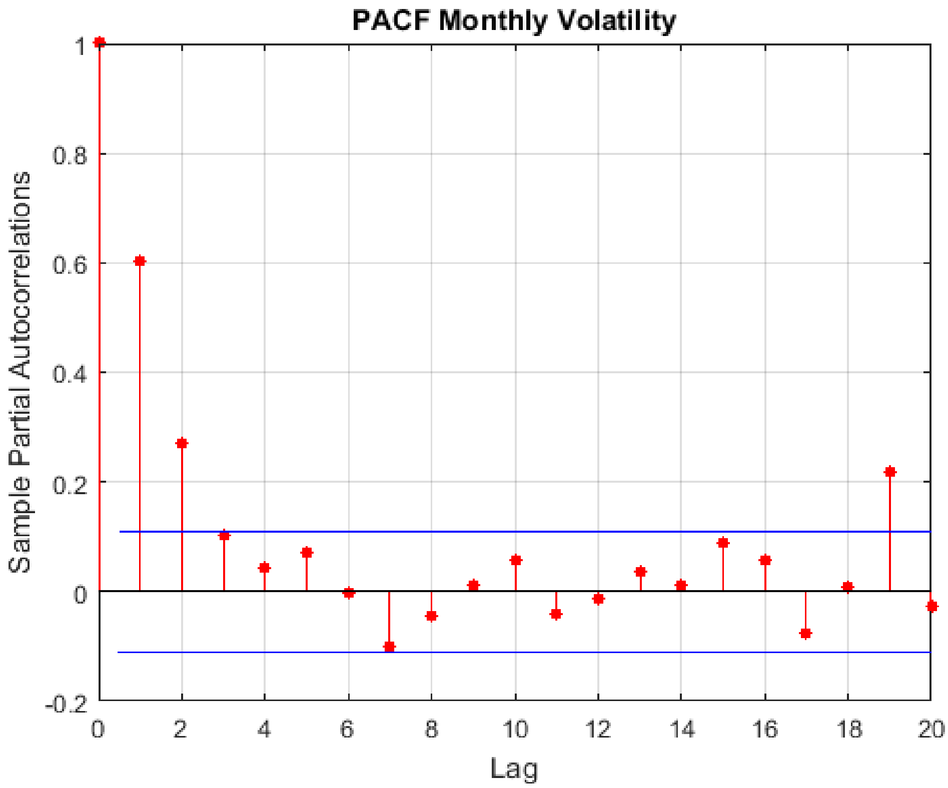


FIGURE B.2

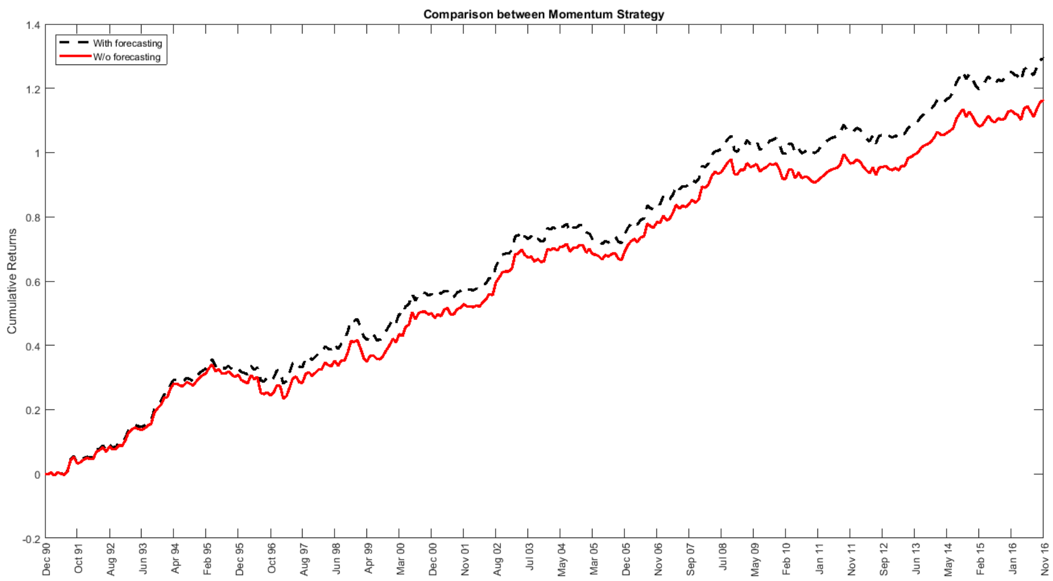


FIGURE B.3

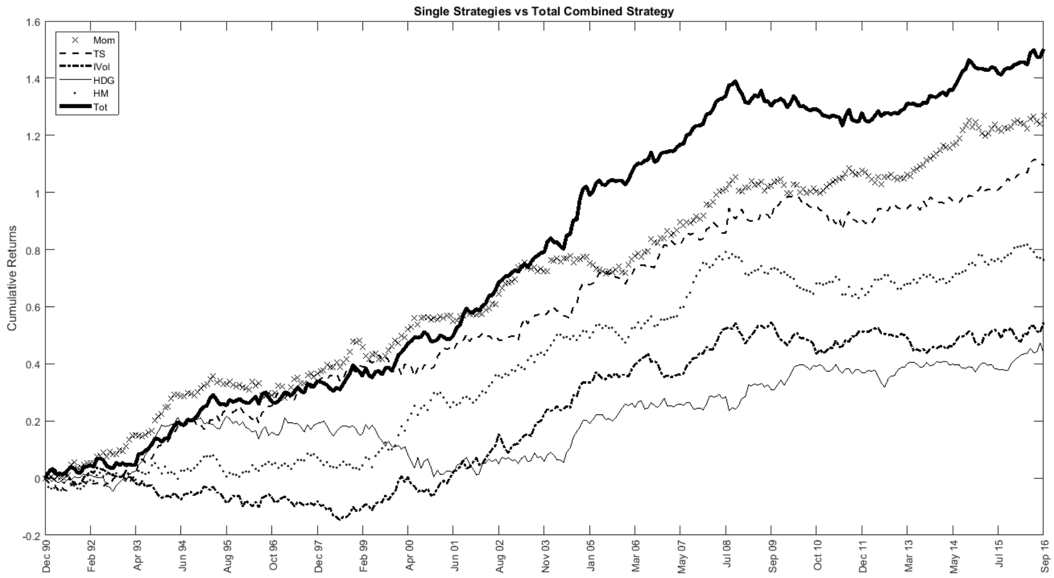


FIGURE B.4

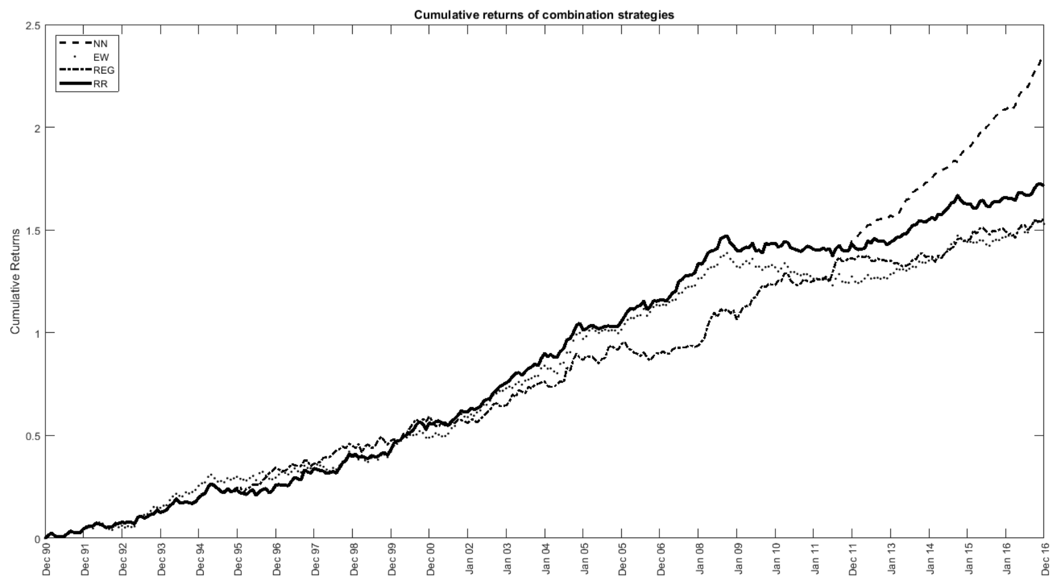


FIGURE B.5

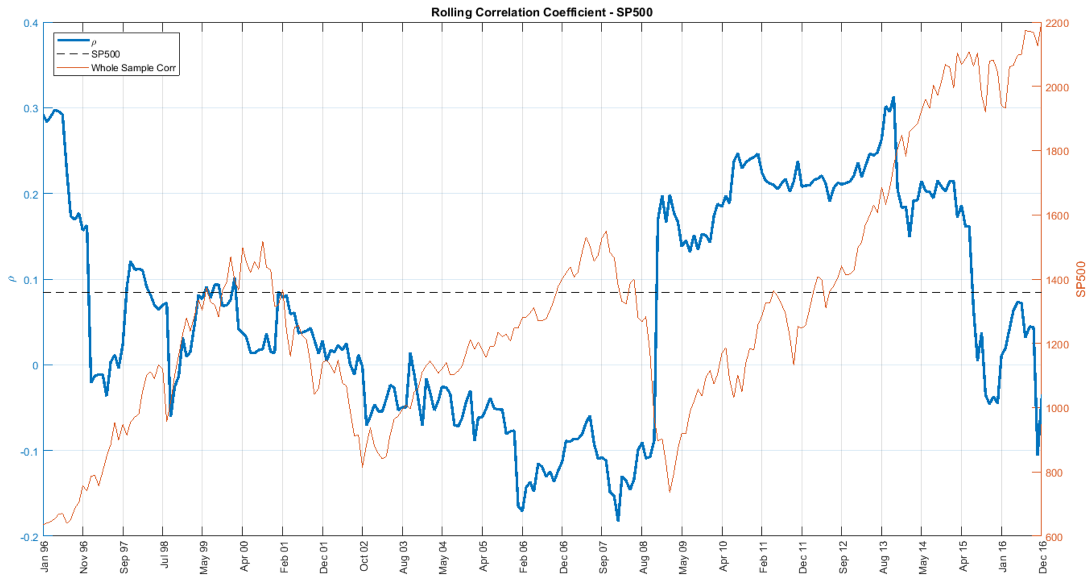

Rolling Correlation Coefficient - SP500
ρ
SP500
Whole Sample Corr
0.4
0.3
0.2
0.1
0
-0.1
-0.2
2200
2000
1800
1600
1400
1200
1000
800
600
SP500
Jan 96
Nov 96
Sep 97
Jul 98
May 99
Apr 00
Feb 01
Dec 01
Oct 02
Aug 03
May 04
Apr 05
Feb 06
Dec 06
Sep 07
Aug 08
May 09
Apr 10
Feb 11
Dec 11
Sep 12
Aug 13
May 14
Apr 15
Jan 16
Dec 16

FIGURE B.6

C References

Achour, D., Harvey, C. R., Hopkins, G., & Lang, C. (1998). Stock selection in emerging markets: Portfolio strategies for Malaysia, Mexico and South Africa. Emerging Markets Quarterly, 2, 38-91.

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. Journal of Finance, 61, 259-299.

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

Bali, T. G., & Cakici N. (2008). Idiosyncratic volatility and the cross-section of expected returns. Journal of Financial and Quantitative Analysis, 43, 29-58.

Bakshi, G., Gao, X., & Rossi, A. (2013). Asset pricing models that explain the cross-section and time-series of commodity returns. Working paper, University of Maryland.

Basu, D., & Miffre, J. (2013). Capturing the risk premium of commodity futures: The role of hedging pressure. Journal of Banking and Finance, 37, 2652-2664.

Barberis, N, Schleifer, A., & Vishny, R. (1998). A model of investor sentiment. Journal of Financial Economics, 49, 307-343.

Cheng, I-H., & Xiong, W. (2013). The financialization of commodity markets. NBER Working paper No. 19642, http://www.nber.org/papers/w19642.

Cooper, M. J., Gutierrez, R. C., & Hameed, A. (2004). Market states and momentum. Journal of Finance, 59, 1345-1366.

Cootner, P. (1960). Returns to speculators: Telser vs. Keynes. Journal of Political Economy, 68, 396-404.

Corsi,F (2009). ”A Simple Approximate Long Memory Model of Realized Volatility” (JFEC)

Daniel, K, Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. Journal of Finance, 53, 1839-1885.

Daskalaki, C., Kostakis, A., & Skiadopoulos, G. (2012). Are there common factors in commodity futures returns? Working Paper, University of Piraeus.

Ding, Z., C. W. J. Granger, and R. F. Engle (1993): “A Long Memory Property of Stock Market Returns and a New Model,” Journal of Empirical Finance, 1, 83–106.

Engle, R. (1982): “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation,” Econometrica, 50, 987–1007.

Engle, R., and T. Bollerslev (1986): “Modeling the Persistence of Conditional Variances,” Econometric Reviews, 5, 1–50.

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

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

Fama, E., & MacBeth, J. D. (1973). Risk, returns, and equilibrium: Empirical tests. Journal of Political Economy, 81, 607-636.

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

Glosten, L., R. Jagannathan, and D. Runkle (1993): "On the relation between expected return on stocks," Journal of Finance, 48, 1779–1801.

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

Goyal, A., & Santa-Clara, P. (2003). Idiosyncratic risk matters! Journal of Finance, 58, 975- 1007.

Granger, C. W. J. (1969). Investigating causal relations by econometric models. Econometrica, 37, 424-438.

Han, Y., & Lesmond, D. (2011). Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. Review of Financial Studies, 24, 1590-1629.

Hong, H., & Stein, J. (1999). A unified theory of underreaction, momentum trading and overreaction in asset markets. Journal of Finance, 54, 2143-2184.

Huang, W., Liu, Q., Rhee, S. G., & Zhang, L. (2010). Return reversals, idiosyncratic risk, and expected returns. Review of Financial Studies, 23, 147-168.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. Journal of Finance, 48, 65-91.

Lambert, P., and S. Laurent (2000): : "Modelling Financial Time Series Using GARCH-Type Models and a Skewed Student Density," Mimeo, Universite de Liege.

Locke, P., & Venkatesh, P. (1997). Futures market transaction costs. Journal of Futures Markets, 17, 229-245.

Merton, R.C. (1973). An intertemporal capital asset pricing model. Econometrica, 41, 867- 887.

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

Miffre, J., Fuertes, A.M., & Fernandez-Perez, A. (2015). Commodity Strategies Based on Momentum, Term Structure and Idiosyncratic Volatility. Working paper.

McNelis, Paul D. Neural Networks in Finance: Gaining Predictive Edge in the Market. Amsterdam: Elsevier, 2007. Print.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. Journal of Finance, 19, 425-442.

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

Stoll, H. & Whaley, R. (2010). Commodity index investing and commodity futures prices. Journal of Applied Finance, 20, 7-46

Working, H. (1949). The theory of the price of storage. American Economic Review, 39, 1254-1266.