Unsupervised Stochastic Arbitrage Discovery: Application in the Commodity Futures Market

Valeria Fedyk[1], Darcy Pu[1] and Hogan Tong[2]
[1]Department of Finance PhD, London Business School
[2]MBA, London Business School

1 Main Idea

We propose a quantitative rules-based strategy, which we term Unsupervised Stochastic Arbitrage Discovery, that trades combinations of cointegrated commodity futures and capitalizes during temporary deviations from the stationarity of their relationship. Specifically, the strategy first identifies linear combinations of commodity futures that form a stationary price series; we call such a stationary price series the spread of a given combination. Whenever this combination spread deviates from its computed fair value, we open a position (going long and short the appropriate futures as dictated by the linear combination coefficients) betting on a reversal back to the spread's fair value. We close the position and realize a profit once the spread reverses.

2 Economic Rationale

The economic rationale behind our strategy is that once a stationary series has been identified, we can use methods such as the Kalman filter to estimate its fair value based on observed data at every point in time and enter into positions whenever the actual observed value of the combination spread drifts significantly far away from the calculated fair value. In a sense, our strategy is an enhancement of a traditional pairs trading strategy where arbitrage opportunities are discovered based on the difference between the observed and filtered prices of an underlying stationary stochastic spread price process, and long and short positions in commodity futures are taken in a predetermined ratio. Indeed as Nicholas (2000) notes, the idea of pairs trading can be applied to any equilibrium relationship in the financial markets.

An underpinning of this strategy is our finding that certain commodity futures appear to be cointegrated, that is to say that there exists a linear combination of them that is stationary. In order to identify such combinations, we first regress the price series of one of the commodities in the proposed combination on the price series of the other commodities in that combination, and then apply the Augmented Dickey-Fuller test on the residuals from this regression. To reduce overfitting, we also use a Ridge regression and cross-validation in the combination selection process, as explained in more detail in Section 3. The number of such combinations that we identify - 61 - is significantly higher than what we would expect to find if none of the combinations spreads were truly stationary.

It is also important to note that several common factors - such as changes in storage costs, commercial hedging demand, business cycles, and inventory decisions - drive a large portion of the changes in commodity futures prices, and can lead commodity futures in a variety of categories to have related economic price determinants which would help explain the source of the cointegration. Within subsets of commodity futures the drivers are even more similar, such as farm yields for agriculture futures and oil drilling volumes for energy futures. Indeed a recent study by Diebold and Yilmaz (2017) finds significant connectedness between commodity futures prices, and reveals not only clear clustering of commodities into groups that match traditional industries but also other connectedness examples, such as the energy sector being most important in terms of sending shocks to other sectors and energy, industrial metals, and precious metals futures being especially tightly connected.

Once a stationary combination spread has been identified, we apply the Kalman filter to the observed spread prices to estimate its fair value as proposed in Elliott and Malcolm (2005). Our trading strategy systematically opens a position whenever the actual combination spread price deviates from its Kalman filter based fair price past the 95th quantile thresholds. We close out our position either when the combination spread returns back to near its estimated fair value or when our loss exceeds 10% (this occurs very infrequently, but is an important stop loss provision we have included for capital preservation risk management purposes). Our strategy is further able to avoid the following issues that often occur to other mean-reverting trading strategies:

1. Failed to provide a proper proxy of the fair value (we are able to provide an estimate of the fair price using the Kalman filter and observed data)

2. Divergence often does not reverse, forcing the investor to stop loss frequently (over 93% of our combinations are profitable)

3. Reversal takes too long (the average time from opening to closing of a position in our strategy is less than 1 week)

While our proposed strategy is entirely systematic, we do acknowledge the importance of constant oversight and occasional discretionary decisions, such as during intense market turmoil. Therefore in any practical application of our strategy in an investment fund, we would also employ discretionary oversight during periods of extraordinary market moves (for instance, we may put our strategy on hold in a high-volatility environment or when we are unsure if a structural break is occurring; in fact we already systematically do so if the combination spread ceases to be stationary in a 150 day lookback period). We would also aim to continuously innovate and improve upon our strategy.

Finally, it is also important to note that the strategy is market neutral and is designed to perform well under any market conditions. As Gorton and Rouwenhorst (2006) note, commodity futures are negatively correlated with equity returns and bond returns due in part to their different behavior over the business cycle; as a consequence, commodity futures can perform well during both market upswings and downturns. As noted by Lintner (1983), commodity futures also constitute a valuable addition to a traditional stock / bond portfolio due to their low correlation with those asset classes and ensuing diversification benefits. Furthermore, Burghardt (2010) notes that the broad commodity trading advisor (CTA) index is very highly correlated (.97) with trend following CTAs; as our strategy is market-neutral, it will offer diversification benefits even among a portfolio of other commodity futures strategies.

3 Implementation

In this section we outline the implementation of our strategy and the systematic rules-based approach we take to construct our portfolio at each point in time.

We start with a total of N commodity futures that we can trade on the major exchanges. For the purposes of our backtest presented in Section 6, we use 33 major commodity futures traded on the major exchanges CME, ICE, LSE, NYMEX, CBOT, SHFE, CZCE, and DCE.

Let us denote the settle price of each commodity i at a given time t as $x_{i,t}$ for $i = 1, 2, .., N$. We define a "commodity combination" as a combination of 2, 3, or 4 commodities that we find to be cointegrated, and the "commodity combination spread" as the corresponding price series $\{c_1 * x_{1,t} + c_2 * x_{2,t} + .. + c_r * x_{r,t}\}_{t=1,2,..}$ for the $r \in \{2, 3, 4\}$ commodities and some constants $c_1, c_2, .., c_r$ for which the price series is stationary. Let us call the coefficients $c_1, c_2, .., c_r$ of the commodity combination spread its positions.

We denote the qth such commodity combination that we find as C_q, and define it as:

$$C_q = \{I_q, \Lambda_q\}$$

where $I_q = (i_{q,1}, .., i_{q,r})$ specifies the indices of the $r \in \{2, 3, 4\}$ commodity futures that form the combination, and $\Lambda_q = (\lambda_{q,1}, .., \lambda_{q,r})$ specifies the associated positions of these commodities in the combination. The price vector of the commodity combination at any given time t is denoted as the tuple of the constituent commodity prices $X_{q,t} = (x_{i_{q,1},t}, .., x_{i_{q,r},t})$.

The combination spread of the combination C_q is thereby defined and calculated as the dot product of the positions and historical prices of each constituent commodity in the combination as follows:

$$y_{q,t} = \Lambda_q X^{'}_{q,t} = \lambda_{q,1} x_{i_{q,1},t} + ... + \lambda_{q,r} x_{i_{q,r},t}$$

For simplicity, we use $y_{q|\Phi}$ to denote $(y_{q,t_1}, y_{q,t_2}, \ldots, y_{q,t_n})'$ for $t_i \in \Phi$. We aim to profit from the mean-reverting movements of $y_{q,t}$, i.e. if $y_{q,t}$ diverges from its fair value, $\bar{y}_{q,t}$ and exceeds a pre-determined threshold, we enter the reverse position.

We propose our strategy, *Unsupervised Stochastic Arbitrage Discovery*, to provide an automated trading system that generates significant returns with manageable drawdowns. In the following sections we detail (i) the procedure and tests for stationarity we used to identify our commodity combinations; (ii) how we used the Kalman filter to compute the fair value of the combination spreads and the ensuing trading indicators; and (iii) our proposed method of capital allocation among the combinations to maintain a significant level of invested capital at all times.

3.1 Identifying Stationary Combinations

In this subsection, we detail our process for determining which commodity combination spreads are stationary. We illustrate our process through an example for a three commodity combination, but the process is analogous for two and four commodity combinations.

A stochastic process is stationary if the unconditional joint probability distribution does not change over time. Consequently, parameters such as mean and variance also do not change over time and we would have confidence that such a series would converge back towards its mean after a period of divergence. If a time series is not stationary, however, there is no guarantee that it will ever converge back after a divergence. Figure 1 depicts the rationale of this aspect of a stationarity price process.

Statistically, we can use the Augmented Dickey–Fuller test (ADF) to determine if a time series is stationary. The Augmented Dickey–Fuller (ADF) statistic, used in the test, is a negative number and a p-value can be deduced from its probability distribution. A p-value smaller than 0.05 indicates that the series is stationary with a confidence level of 95%.

To generate a candidate combination q, we first randomly select three commodities, $(i_{q,1}, i_{q,2}, i_{q,3})$, out of a total of $\binom{N}{3}$ possible combinations and run the following regression:

$$x_{i_{q,1},t} = a + b x_{i_{q,2},t} + c x_{i_{q,3},t} + \epsilon_{q,t}$$

We then test the stationarity of the tradable combination spread $\{x_{i_{q,1},t} - b x_{i_{q,2},t} - c x_{i_{q,3},t}\}$ which is equivalent to testing the stationarity of $\epsilon_{q,t}$. If the p-value of the Augmented Dickey-Fuller test is $\leq 5\%$, then the combination $\{(i_{q,1}, i_{q,2}, i_{q,3}), (1, -b, -c)\}$ is selected. The intuition is that if these three commodities indeed have a stable relationship, then the residuals of the above regression should be independently distributed, and thus $y_{q,t} = x_{i_{q,1},t} - b x_{i_{q,2},t} - c x_{i_{q,3},t}$ should be stationary.

We further extended the above approach by using Ridge regression and cross validation to avoid over-fitting, i.e. the situation when $y_{q,t}$ is stationary in the training sample, but non-stationary in the test sample due to fitting the training sample data too closely. In order to combat overfitting, we divide the training sample into four periods of equal length, Φ_1, Φ_2, Φ_3 and Φ_4, and then run Ridge regressions on every three-commodity combinations using only the data in Φ_1. We further test the robustness of our model by varying the penalty λ in the Ridge regression, and find that our results do not change significantly with λ varying from $e^{-1} \leq \lambda \leq 0$ (also note that the standard OLS regression is a special case of Ridge regression corresponding to a λ of 0). The Ridge regression we run is:

$$x_{i_{q,1},t} = a + b x_{i_{q,2},t} + c x_{i_{q,3},t} + \epsilon_{q,t} \text{ for } t \in \Phi_1$$

The combination $\{(i_{q,1}, i_{q,2}, i_{q,3}), (1, -b, -c)\}$ is selected if $y_{q,t} = x_{i_{q,1},t} - bx_{i_{q,2},t} - cx_{i_{q,3},t}$ passes the stationarity test in each time period Φ_1, Φ_2, Φ_3 and Φ_4. In other words, we derive the positions Λ_q using data in Φ_1, but we keep C_q only if it passes stationarity tests in all Φ_1, Φ_2, Φ_3 and Φ_4. This cross-validation step, combined with using a Ridge regression, greatly reduces the probability of overfitting.

3.2 Calculating Fair Value and Trading Thresholds

On each trading day t, we consider a $l = 150$ day lookback period to (i) determine that the series has remained stationary, (ii) compute fair value, and (iii) determine our trading thresholds. Let $y_{q,t,l}$ denote the time series of $y_{q,t}$ in $(t - l, t)$. For each combination C_q at each time t, we calculate the following:

1. The **fair value** of $y_{q,t}$ is calculated by applying the Kalman filter[1] to $y_{q,t,l}$ and the fair value is the priori state estimate of $y_{q,t}$. Furthermore, we apply the Shumway & Stoffer smoother and EM algorithm in estimating the Kalman filter parameters as outlined by Shumway and Stoffer (1982). We denote the fair value of $y_{q,t}$ as $\bar{y}_{q,t}$.

2. The **trading threshold** is the 95% quantile of $|y_{q,t,l} - \bar{y}_{q,t,l}|$, i.e. the distance between observed value and fair value. The trading threshold is denoted as $\theta_{q,t}$

3. The **close threshold** is the 50% quantile of $|y_{q,t,l} - \bar{y}_{q,t,l}|$, denoted by $\delta_{q,t}$

The trading strategy for the combination C_q can be summarized in pseudocode according to Algorithm 1. In short, we enter the reverse position if $y_{q,t}$ is outside the band of $\bar{y}_{q,t} \pm \theta_{q,t}$, and close the position if $y_{q,t}$ converges to the band of $\bar{y}_{q,t} \pm \delta_{q,t}$.

Furthermore, we add a loss control measure: in any given time point t, if daily loss exceeds 10%, we close the position and enter a 10-day cool down period, i.e. do not trade for ten days no matter what.

3.3 Allocating Fund Capital

We have established how to construct a universe of commodity combinations and the trading strategy of each combination. We now propose two capital allocation strategies, first-come-first-serve allocation and rank-based weighting allocation.

First-Come-First-Serve Allocation (FCFS): At any time t, if a combination shows a trading signal, we allocate all the available capital to that combination. If more than one combination shows a trading signal, we allocate available capital evenly among them. If a combination shows a trading signal but there is no available capital, we do not trade it. When a combination in trading shows close signal, we close the position and free up the capital.

Rank-Based Weighting Allocation (RBW): We rank combinations based on their performance in the 150 trading days before the testing period. A higher rank indicates higher cumulative return in those 150 trading days. We denote the rank of each C_q as R_q, and its position at time t as $pos_{q,t} \in \{-1, 0, 1\}$. At any time t, the weight of C_q in a portfolio with N combinations is:

$$w_{q,t} = \frac{|pos_{q,t}|R_q}{\sum_{q=1}^{N} |pos_{q,t}|R_q}$$

In other words, combinations with higher cumulative returns in the training sample will be assigned higher weights in the testing sample.

[1] More on Kalman filter: https://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/

4 Risk Management

In constructing our strategy, we have kept risk management considerations at the forefront of our development goals. Having not only academic but also practical fund management experience, we know all too well the benefits of avoiding overconfidence and allocating capital to strategies with stop loss provisions and position limits, as well as the dangers of not doing so.

Our risk management strategy is centered on three main objectives:

1. **Preserve client credibility:** First and foremost, we aim to always maintain our clients' trust in our investment strategy. For this reason, it is imperative that we do our best to limit large sudden losses, as such losses would risk eroding investor trust in our fund. We implement a 10% stop loss provision, where if daily losses in a combination ever exceed 10% of the NAV allocated to that combination, we liquidate those positions and enter a 10 day cool down period where we do not trade in the combination.

2. **Maintain adequate investable capital:** Our second objective is to maintain a steady level of investable capital at all times. We wish to avoid large losses or similar events that would wipe out a sufficient amount of the fund capital to never be able to fully recover to pre-loss levels. For instance, if a fund's NAV falls by 70% one period, it would need to rise by 233% (an order of magnitude higher) in order to return to the initial NAV level, a very difficult return level to achieve. Compounding on this difficulty is the fact that large sudden losses or lengthy periods of bad performance often lead investors to redeem their investments rather than waiting for the strategy to recover, which comes at the worst possible time for the strategy. Our stop loss measure is one measure towards preventing this situation from arising, and we would also aim to educate our clients on the dangers of withdrawing capital during downturns as outlined by Ghemawat (1993). Finally, commodity futures investment often involve a significant amount of leverage, which raises the probability of a forced and undesirable margin call, which our strategy avoids by maintaining a 20% margin at all times and closing out the position when the loss reaches this margin amount.

3. **Preserve strategy performance:** A third risk we aim to protect against comes from market conditions under which our strategy ceases to perform well, namely when a structural change causes a combination spread to cease being stationary. As outlined in Section 2, our strategy is predicated on certain combinations of commodity futures having a stationary relationship through time barring any large, significant structural changes. Consequently, if a structural change does occur, we would wish to quickly close out our positions and put a hold on our strategy while we reassess the market environment. For this reason, we implement a 150 day lookback period which verifies that the combination spread has remained stationary over that period before entering any positions. This 150 day lookback period also allows us to recalibrate the quantile threshold at which we enter into positions to the most recent and relevant data available (for example, if the volatility of the spread decreases, our 95% quantile thresholds would correspond to lower absolute value changes, making it easier for us to enter into a position).

We hope that having these risk management measures in place - including stop loss provision, lookback period, and client advisory services during downturns - would allow us to build a strategy with stable positive performance over time, retain and build trustworthy relationships with clients, and lead to long-term profitability.

5 Liquidity, Tax and Capital Considerations

Our strategy trades 33 commodity futures from the major exchanges CME, ICE, LSE, NYMEX, CBOT, SHFE, CZCE, and DCE, which are highly liquid and regulated instruments. As Abrams and Flores (2014) note, the futures market has grown in the past decades and this growth has contributed to increased depth and liquidity in the futures markets. Consequently we do not anticipate illiquidity concerns in our strategy, which is a benefit as we would be able to scale up our strategy more easily. Specifically, our strategy trades on the order of 1000 commodity contracts per day while the total traded volume of commodity futures contracts in 2019 according to a study by WFE (2019) was 6,569 million contracts, or 28 million contracts per trading day. Consequently, our average daily trades currently constitute a miniscule fraction (less than 0.0001%) of the total daily contracts traded, leaving ample room for us to scale up without affecting undesired price impact with our trades.

The liquidity of our strategy, as well as the fact that our average holding period for a position is less than one week, would also allow us to return capital to investors who wish to redeem under a relatively short notice period without necessitating a detraction from our strategy by prematurely closing out our positions. Not requiring a stringent capital lock-up periods for investors would also be an attractive proposition of our strategy when marketing it to clients.

We anticipate our investment vehicle to be structured as a Commodity Trading Advisor (CTA) fund, or a hedge fund that uses futures contracts to achieve its investment objective. In terms of capital considerations, for every commodity futures contract that we purchase we would post a 20% margin, a much higher ratio than current market requirements. We close out our position if the loss ever reaches 10% of the notional amount, which limits our risk of a sudden margin call.

With regards to tax, according to the Tax Act of 1981 short-term profits in futures are taxed at 60% long-term (therefore subject to a maximum of 15%) and 40% short-term (normal taxable income) rate. This is in contrast to equity trading, where 100% of short-term profits are taxed at the higher short-term rate, so our investment vehicle is also advantageous from a tax perspective.

6 Analysis of Strategy Prospects

We now turn to looking at the hypothetical performance of our strategy in order to analyze the expected risks and returns. Our backtest uses commodity pricing data obtained from WindPy over the period 02/13/2012 - 01/22/2021, which constitutes 2,251 trading days or roughly 9 years of trading data. The dataset includes 33 commodity futures traded on the major exchanges CME, ICE, LSE, NYMEX, CBOT, SHFE, CZCE, and DCE.

In order to identify the stationary commodity combination spreads and retrieve the corresponding position values for such combinations, we use the approach outlined in Section 3.1 and pricing data from the first 1200 days of our dataset as our training sample. We then use the identified combinations and their parameters to run a backtest of our strategy over the 1000 trading days over 02/02/2017 - 01/22/2021. Note that the first 150 days correspond to the lookback period, so the backtest strategy returns are computed from 09/07/2017 - 01/22/2021.

We find that our trading strategy generates an annualized return of 88.91% under the First-Come-First-Serve (FCFS) allocation method and an annualized return of 90.01% under the Rank-Based Weighting (RBW) allocation method. Capital utilization is well below limit at 60.75% and 73.18% for the FCFS and RBW capital allocation methods respectively, and the maximum drawdowns are at a high but manageable 28.22% and 38.87%. The Sharpe ratio for the trading strategy in the backtest under both allocation methods nears 2. The resulting backtest statistics are summarized in Table 1 and growth of portfolio NAV under two capital allocation methods are shown in Figure 3.

Bibliography

Abrams, R., B. R. and E. Flores (2014): "A Quantitative Analysis of Managed Futures Strategies," *Lintner Revisited*.

Burghardt, G. (2010): "Measuring the impact of trend following in the CTA space," *Opalesque TV*.

Diebold, F. X., L. L. and K. Yilmaz (2017): "Commodity Connectedness," *Working Paper*.

Elliott, R. J., v. d. H. J. and W. P. Malcolm (2005): "Pairs Trading," *Quantitative Finance 5 (3): 271-276*.

Ghemawat, P. (1993): "The Risk of Not Investing in a Recession," *MIT Sloan Management Review 34 (2): 51*.

Gorton, G. and J. G. Rouwenhorst (2006): "Facts and Fantasies about Commodity Futures," *Financial Analysts Journal 62 (2): 47-68*.

Lintner, J. V. (1983): "The Potential Role of Managed Commodity-financial Futures Accounts (and/or Funds) in Portfolios of Stocks and Bonds," *Issue 79 of Working paper (Harvard University. Graduate School of Business Administration. Division of Research)*.

Nicholas, J. G. (2000): *Market Neutral Investing: Long / Short Hedge Fund Strategies*, Bloomberg Professional Library, Bloomberg Press: Princeton, NJ, USA.

Shumway, R. H. and D. S. Stoffer (1982): "An approach to time series smoothing and forecasting using the EM algorithm," *Journal of Time Series 3 (4): 253-264*.

WFE (2019): "The WFE's Derivatives Report 2019," *WFE Research*.

Appendix A. Implementation Detail

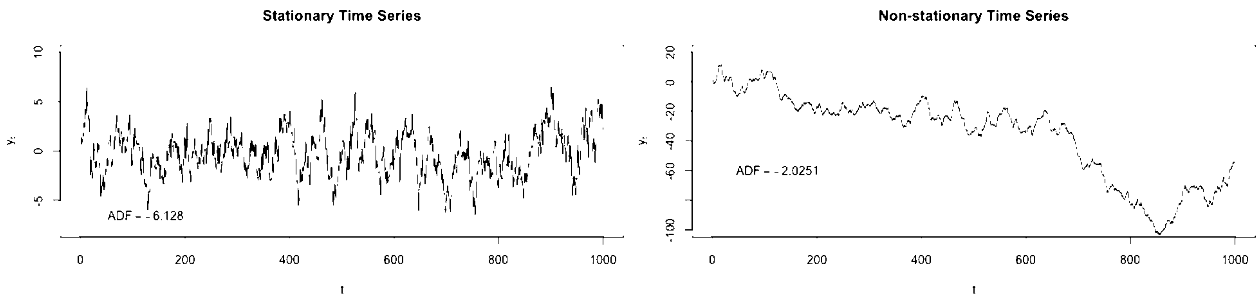


Figure 1: Stationary vs. Non-Stationary Time Series. The left chart is stationary and the right chart is non-stationary. A mean-reverting trading strategy would only work with the left chart.

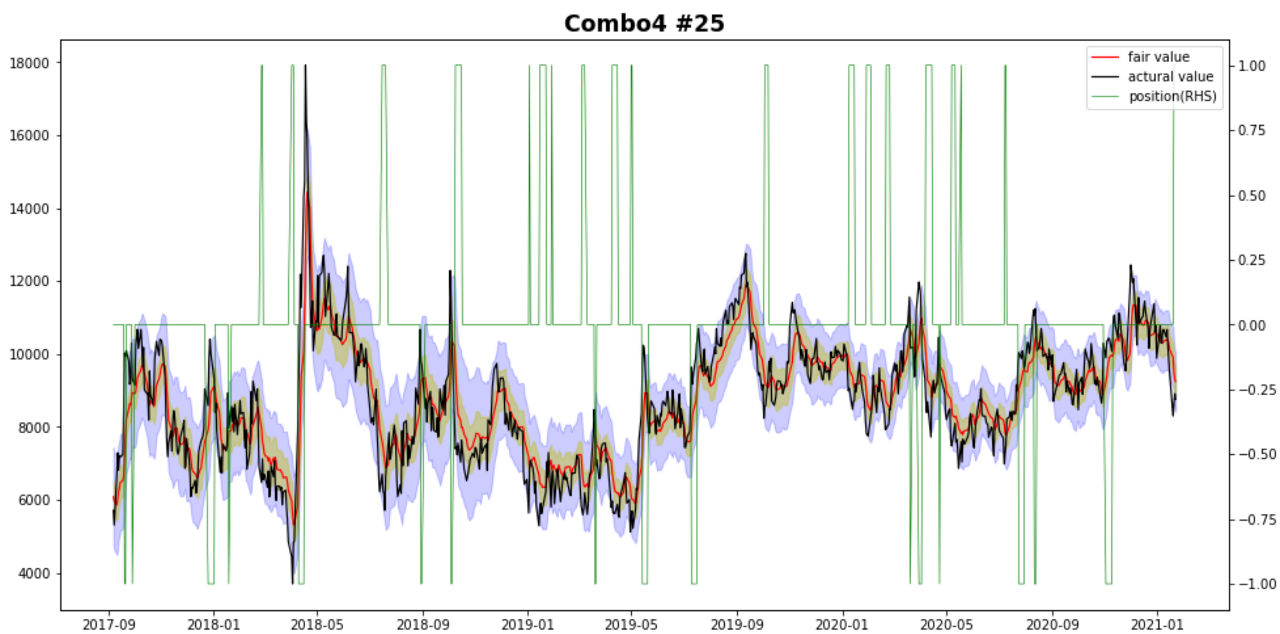


Figure 2: Sample 4-commodity combination. The black line represents the actual value of the commodity combination spread over time, the red line represents the fair value as calculated using the Kalman filter for the commodity combination spread over time, and the green line represents the long and short positions taken in the combination over time. The blue band represents the trading threshold and the yellow band represents the close threshold. Note that actual value being significantly less than the computed fair value coincides with buying the combination next period, while actual value being significantly more than the computed fair value coincides with selling the combination next period.

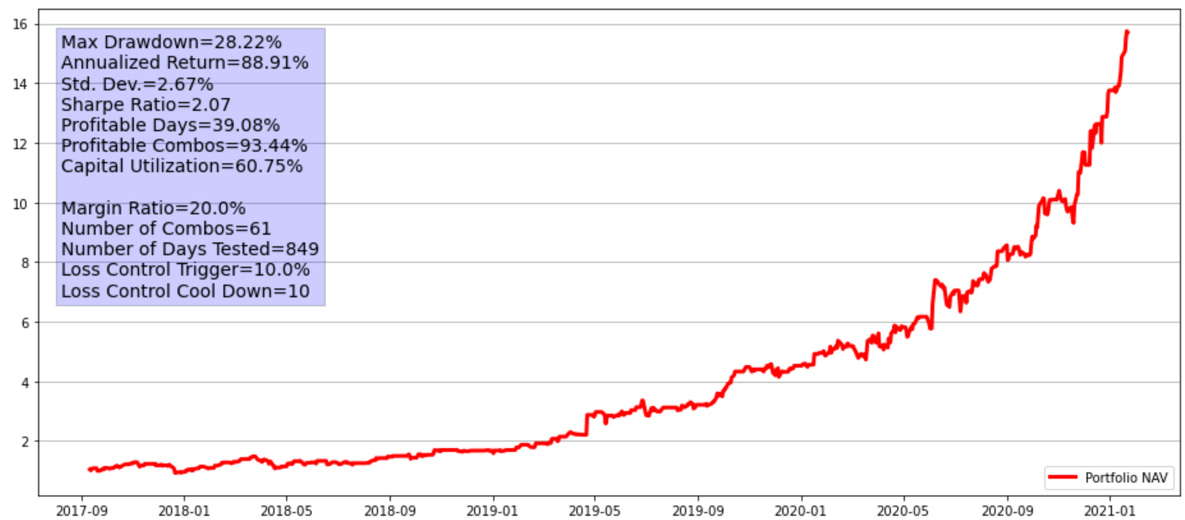
Algorithm 1 Unsupervised Stochastic Arbitrage Discovery

for each $t \in \Phi_{test}$ **do**
  if $y_{q,t,1}$ fails stationarity test **then**
    $position_{q,t} \leftarrow 0$ and **continue loop** ▷ If C_q is non-stationary, close position
  else if $position_{q,t-1} = 0$ **then** ▷ Currently not holding a position in C_q
    if $y_{q,t} > \bar{y}_{q,t} + \theta_{q,t}$ **then** ▷ Observed spread is above upper trading threshold
      $position_{q,t} \leftarrow -1$ ▷ Short C_q
    else if $y_{q,t} < \bar{y}_{q,t} - \theta_{q,t}$ **then** ▷ Observed spread is below lower trading threshold
      $position_{q,t} \leftarrow 1$ ▷ Long C_q
    else ▷ Observed spread does not trigger a trade
      $position_{q,t} \leftarrow 0$ ▷ Do nothing
    end if
  else if $position_{q,t-1} = -1$ **then** ▷ Currently holding a short position in C_q
    if $y_{q,t} < \bar{y}_{q,t} + \delta_{q,t}$ **then** ▷ Observed spread is below upper close threshold
      $position_{q,t} \leftarrow 0$ ▷ Close the short position in C_q
    else ▷ Observed spread is above upper close threshold
      $position_{q,t} \leftarrow -1$ ▷ Keep shorting C_q
    end if
  else if $position_{q,t-1} = 1$ **then** ▷ Currently holding a long position in C_q
    if $y_{q,t} > \bar{y}_{q,t} - \delta_{q,t}$ **then** ▷ Observed spread is above lower close threshold
      $position_{q,t} \leftarrow 0$ ▷ Close the long position in C_q
    else ▷ Observed spread is below lower close threshold
      $position_{q,t} \leftarrow 1$ ▷ Keep longing C_q
    end if
  end if
end for each

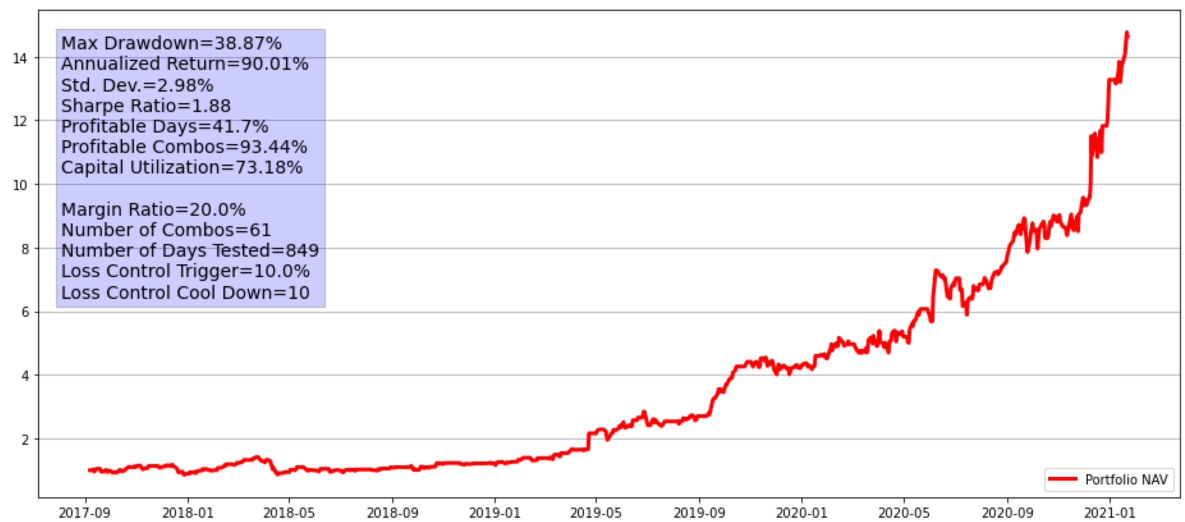
Appendix B. Backtest Performance

Allocation Method	Annualized Return	Sharpe Ratio	Capital Utilization	Max Drawdown	Profitable Time
FCFS	88.91%	2.06	60.75%	28.22%	39.08%
RBW	90.01%	1.88	73.18%	38.87%	41.70%

Table 1: Risk and return statistics for the trading strategy backtest over 09/2017 - 01/2021.



(a) First-Come-First-Serve (FCFS) capital allocation method



(b) Rank-Based Weighting (RBW) capital allocation method

Figure 3: Growth of NAV for the trading strategy backtest over September 2017 - January 2021 using both capital allocation methods. NAV is normalized to 1 at the beginning of period.