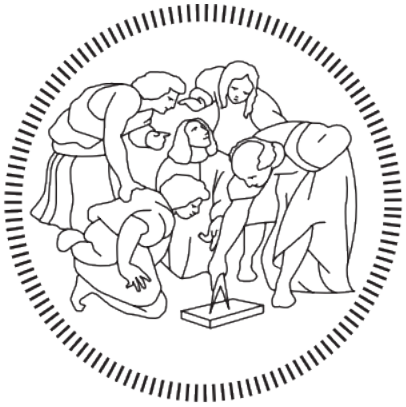


POLITECNICO
MILANO 1863

Financial Engineering

Final Project Group 8B

2024-2025

Name	CP	MAT
Giacomo Kirn	10773720	275759
Marco Amarilli	10729094	260400
Ismail Oulkhir	11077020	277280

Contents

0. Introduction

The goal of this paper is to analyze the performance and robustness of a **statistical arbitrage strategy** based on a **mean-reverting spread** in the energy futures market. Specifically, we study the pair composed of Heating Oil (`HOc2`) and Gas-Oil (`LGOc6`) over the period from April 22, 2015 to April 21, 2016, using high-frequency data sampled every 30 minutes.

Following the framework proposed by **Baviera and Santagostino** (2019), the strategy is developed under the assumption that the log-ratio of the prices follows an **Ornstein–Uhlenbeck (OU)** process, and allows for the use of both **leverage** and a **stop-loss mechanism**. The long-run return of the strategy can be expressed analytically as a function of the OU parameters, transaction costs, trading thresholds, and leverage factor.

The structure of this project is as follows:

- In **Section 1**, we provide a theoretical proof that a profitable strategy can be constructed under the OU model, even in the presence of a stop-loss, provided that transaction costs remain below a critical level. This includes a reproduction of **Figure 3** from the original paper.
- **Sections 2–5** are devoted to the empirical calibration and performance analysis of the strategy on In-Sample and Out-of-Sample data. We estimate the OU parameters, assess the impact of different time windows, and study the sensitivity to various stop-loss thresholds.
- **Section 6** introduces a dedicated framework for computing **confidence intervals** on key performance metrics using a parametric bootstrap combined with interpolation over a precomputed optimization grid. We address discontinuities in optimal leverage and discuss statistical treatment for samples in which the strategy becomes unprofitable.
- **Section 7** We introduce a conceptual extension of the evaluation framework to account for **parameter uncertainty** through a bootstrapped out-of-sample simulation procedure. By generating multiple strategy configurations via resampling of in-sample data and applying each to a fixed test set, this approach enables a distributional assessment of portfolio outcomes. It allows for the computation of **risk metrics** such as return confidence intervals, probability of ruin, and drawdowns, thereby providing deeper insights into the strategy's robustness under model uncertainty, trading band variation, and leverage.
- **Sections 8–9** extend the application of the model to alternative asset classes, including governative bond futures and cryptocurrency pairs, with an emphasis on testing the OU assumption and adapting the strategy across markets.
- Finally, **Section 10** concludes the report with a critical assessment of the strategy's strengths and weaknesses, offering practical recommendations for real-world implementation.

In addition to the report, we provide two complete codebases developed in:

- **MATLAB**, which contains all answers to the mandatory questions, as well as the solution to the first optional question (Question f), which analyzes optimal statistical arbitrage with stop-loss and leverage on governative futures;
- **Python**, which also covers all the mandatory questions and provides the solution to the second optional question (Question g), focused on cryptocurrency pair trading.

Throughout the analysis, we make extensive use of the analytical methods and numerical techniques introduced in Baviera and Santagostino 2019 to assess both theoretical soundness and empirical performance under realistic market conditions.

1. Question a

Context. We consider a statistical arbitrage strategy based on repeated long/short positions in a mean-reverting asset, where trades are opened at a fixed entry level and closed at either a take-profit or stop-loss threshold. The position is subject to transaction costs. This framework, formalized in Baviera and Santagostino (2019), allows for the analytical characterization of profitability conditions under general stochastic dynamics, with closed-form expressions available in the case of Ornstein–Uhlenbeck (OU) processes.

Long-run return. Let X_t denote the process driving the strategy (e.g., the log-price ratio between two co-integrated assets). A trade is opened at level D, and closed at either $U > D$ (take-profit) or $L < D$ (stop-loss), incurring transaction cost $C > 0$. The long-run return of the strategy under leverage f is given in general form by:

$$\mu(f, D, U \mid L) = \frac{p^+ \ln(1 + fv^+) + p^- \ln(1 + fv^-)}{p^+ \mathbb{E}[t^+] + p^- \mathbb{E}[t^-]}, \tag{1}$$

where p^+ is the probability of exit at U, p^- at L, $v^\pm$ are the net returns after cost, and $\mathbb{E}[t^\pm]$ are the conditional expected durations of the trades.
In the case where X_t follows an Ornstein–Uhlenbeck process with mean-reversion speed k, volatility σ, and stationary standard deviation $\Sigma = \sigma/\sqrt{2k}$, it is convenient to introduce the following rescaled variables:

$$d := \frac{D}{\Sigma}, \quad u := \frac{U}{\Sigma}, \quad \ell := \frac{L}{\Sigma}, \quad c := \frac{C}{\Sigma}.$$

This transformation expresses the entry/exit/stop-loss levels and transaction cost in units of Σ.
In this framework, the long-run return takes the following closed-form expression:

$$\mu(f, D, U \mid L) = \frac{1}{\pi\theta} \left(\frac{\ln\left[1 + f\left(e^{U-D-C} - 1\right)\right]}{\mathrm{Erfid}(u, d)} + \frac{\ln\left[1 + f\left(e^{L-D-C} - 1\right)\right]}{\mathrm{Erfid}(d, \ell)} \right), \tag{23}$$

where $\theta = 1/k$ is the characteristic time scale of the OU process.

Profitability condition. We now formalize the conditions under which a statistical arbitrage strategy based on an Ornstein–Uhlenbeck process is profitable, both in absolute terms and in the low-volatility regime.

Proposition 1.1 (Critical transaction cost)**.** *Let $L < D < U$ be a fixed stop-loss and trading band pair, and let p^+ denote the probability of hitting the upper band U before the stop-loss L. Then the strategy admits a strictly positive optimal leverage (i.e., is profitable) if and only if the transaction cost C satisfies:*

$$C < \bar{C}(D, U, L) := L - D + \ln\left(1 + p^+\left(e^{U-L} - 1\right)\right). \tag{2}$$

Proof. We begin from the profitability condition (see Lemma 1 in Baviera and Santagostino (2019)):

$$p^+ > q^+, \qquad q^+ := \frac{v^-}{v^- - v^+},$$

with

$$v^+ = e^{U-D-C} - 1, \quad v^- = e^{L-D-C} - 1.$$

Substituting into q^+, we get:

$$q^+ = \frac{e^{L-D-C} - 1}{e^{L-D-C} - e^{U-D-C}}.$$

The inequality $p^+ > q^+$ is then equivalent to:

$$p^+(e^{L-D-C} - e^{U-D-C}) < e^{L-D-C} - 1,$$

since $L < D$ and the denominator is negative.
Rewriting and isolating e^C, we factor both sides:

$$e^C < e^{L-D}\left(1 - p^+(1 - e^{U-L})\right),$$

and finally:

$$C < L - D + \ln\left(1 + p^+(e^{U-L} - 1)\right). \qquad \square$$

Low-volatility approximation. We now explore how this condition simplifies in the common case where the stationary standard deviation $\Sigma = \sigma/\sqrt{2k}$ is small, as typically occurs in spread-based arbitrage.

Proposition 1.2 (Profitability under low volatility)**.** *Let $\ell < d < u$ and define the critical cost in Σ units:*

$$\bar{c}_\ell(d, u) := p^+(u - \ell) - (d - \ell).$$

Then, in the low-volatility regime $\Sigma \to 0$, the strategy is profitable if and only if:

$$c < \bar{c}_\ell(d, u).$$

Proof. We expand the critical cost function $\bar{C}(\Sigma)$ from Proposition 1.1:

$$\bar{C}(\Sigma) = \ln\left(1 + p^+\left(e^{(u-\ell)\Sigma} - 1\right)\right) - (d - \ell)\Sigma.$$

We compute the first derivative of $\bar{C}(\Sigma)$ with respect to Σ:

$$\frac{d\bar{C}}{d\Sigma} = \frac{p^+(u - \ell)e^{(u-\ell)\Sigma}}{1 + p^+\left(e^{(u-\ell)\Sigma} - 1\right)} - (d - \ell).$$

Evaluating at $\Sigma = 0$, we obtain:

$$\left.\frac{d\bar{C}}{d\Sigma}\right|_{\Sigma=0} = \frac{p^+(u - \ell)}{1} - (d - \ell) = p^+(u - \ell) - (d - \ell) = \bar{c}_\ell(d, u).$$

Therefore, for small Σ, the first-order expansion is:

$$\bar{C}(\Sigma) = \bar{c}_\ell(d, u) \cdot \Sigma + \mathcal{O}(\Sigma^2).$$

As defined above, $C = c \cdot \Sigma$. Substituting into the profitability condition $C < \bar{C}(\Sigma)$, we obtain:

$$c \cdot \Sigma < \bar{c}_\ell(d, u) \cdot \Sigma.$$

Since $\Sigma > 0$, this simplifies to:

$$c < \bar{c}_\ell(d, u).$$

This concludes the proof. $\square$

1.1. Numerical illustration

Reproduction of Figure 3 . To illustrate the existence of a critical transaction cost $c^*(\ell)$ and the associated optimal trading bands, we compute numerically the values of the entry and exit thresholds (d, u) for increasing transaction costs, under a fixed stop-loss level $\ell = -1.960$. As shown in Baviera and Santagostino (2019), the strategy is well defined and profitable for any $c < c^*(\ell)$, and becomes infeasible when $c \geq c^*(\ell)$.

The following observations can be made:

- The optimal bands (d, u) are no longer symmetric due to the presence of the stop-loss constraint;
- As $c \to c^*(\ell)$, the trading bands converge, and the long-run return tends to zero;
- In the low volatility regime (small Σ), the long-run return remains strictly positive for all $c < c^*(\ell)$.

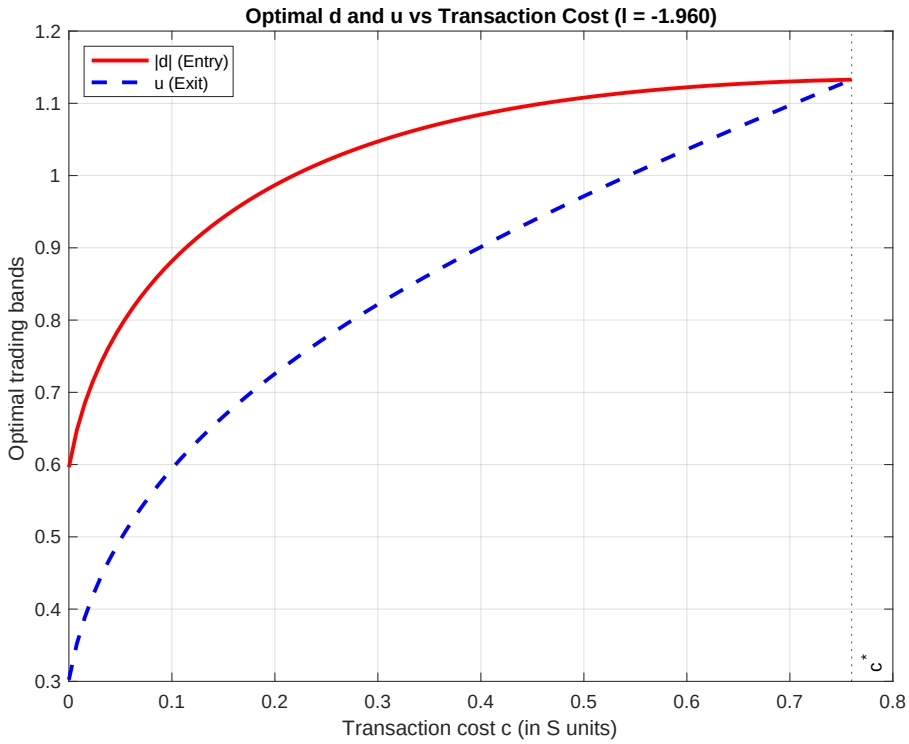


Figure 1: Reproduction of Figure 3 from Baviera and Santagostino (2019): optimal trading bands $|d|$ and u as a function of transaction cost c for a fixed stop-loss $\ell = -1.960$.

2. Question b

2.1. Unit Conversion for HOc2 and LGOc6

Our raw data come from `HO-LGO.xlsm`, which records mid-quotes for two futures contracts with different pricing units:

- **HOc2** (Heating Oil): quoted in \$/gallon.
- **LGOc6** (Gas oil): quoted in \$/metric ton.

To make the two price series directly comparable, we convert both to a consistent unit: dollars per barrel.

Conversion details.

- Since 1 barrel = 42 gallons, we convert HOc2 from \$/gallon to \$/barrel by multiplying by 42:

$$\tilde{p}_{\mathrm{HO},t} = 42 \cdot p_{\mathrm{HO},t}.$$

- According to ICE Futures Europe contract specifications ICE Futures Europe 2024, 100 metric tons of Gas oil correspond to 118.35 cubic meters, implying a standard density of 0.845 kg/liter at 15°C. This yields an approximate conversion of 7.44 barrels per metric ton. Therefore, we convert LGOc6 to \$/barrel by dividing by 7.44:

$$\tilde{p}_{\mathrm{LGO},t} = \frac{1}{7.44} \cdot p_{\mathrm{LGO},t}.$$

With both series now expressed in \$/barrel, the log-ratio of the prices is defined as:

$$R_t = \ln\left(\frac{\tilde{p}_{\mathrm{HO},t}}{\tilde{p}_{\mathrm{LGO},t}}\right) = \ln\left(\frac{42\,p_{\mathrm{HO},t}}{p_{\mathrm{LGO},t}/7.44}\right) = \ln(p_{\mathrm{HO},t}) - \ln(p_{\mathrm{LGO},t}) + \ln(42) + \ln(7.44).$$

Equivalently:

$$R_t = \ln\left(\frac{312.48\,p_{\mathrm{HO},t}}{p_{\mathrm{LGO},t}}\right) = \ln(312.48) + \ln(p_{\mathrm{HO},t}) - \ln(p_{\mathrm{LGO},t}).$$

2.2. Outlier Filtration

Following the methodology described by the paper Baviera and Santagostino 2019, we identify extreme outliers in the log-ratio series R_t. Let Q_1 and Q_3 denote the lower and upper quartiles, respectively, and define the interquartile range as $\mathrm{IQR} := Q_3 - Q_1$. An observation is considered an outlier if it satisfies either of the following conditions:

$$R_t < Q_1 - 3 \times \mathrm{IQR} \quad \text{or} \quad R_t > Q_3 + 3 \times \mathrm{IQR}.$$

Applying this criterion, no extreme outliers are detected in the dataset.
To enhance robustness, we further filter out *antipersistent outliers*—observations that exhibit sharp temporary deviations likely to distort the trading strategy. Specifically, an observation R_t is classified as an antipersistent outlier if:

$$|R_t - R_{t-1}| > \mathrm{IQR} \quad \text{and} \quad |R_{t+1} - R_t| > 0.95 \times \mathrm{IQR}.$$

Using this rule, we remove one antipersistent outlier in the in-sample (IS) period, occurring on 03-Jul-2015 09:00:00 NYT, and one in the out-of-sample (OS) period, occurring on February 15, 2016 at 20:00 NYT.

2.3. Model Calibration

Objective. The goal of this section is to calibrate an Ornstein–Uhlenbeck (OU) process on the log-ratio of futures prices $R_t := \log\left(\frac{p_t^{\text{HOc2}}}{p_t^{\text{LGOc6}}}\right)$, using only business hours data in the In-Sample (IS) period. We estimate the model parameters κ (mean reversion speed), η (long-run equilibrium level), and σ (diffusion parameter), and compute their **bootstrap confidence intervals** following the methodology in Baviera and Santagostino (2019).

We evaluate the sensitivity of the parameter estimates to the intraday time filter, comparing two windows:

- **9:00–16:00 NYT**, the benchmark period suggested in the reference paper;
- **8:00–16:00 NYT**, to assess whether earlier market activity introduces significant bias.

Model and estimation. The log-price ratio R_t is assumed to follow the OU stochastic differential equation:

$$dX_t = \kappa(\eta - X_t)dt + \sigma dB_t$$

where:

- $\kappa > 0$ is the mean-reversion rate;
- η is the long-run mean;
- σ is the instantaneous volatility;
- B_t is a standard Brownian motion.

Let $(x_0, x_1, \ldots, x_N)$ be the observed values of R_t sampled at time intervals of size Δt (in years). The **discrete-time transition** of the OU process is normally distributed:

$$X_{t+\Delta t} \mid X_t \sim \mathcal{N}\left(\mu_t,\, \nu_t^2\right)$$

with:

$$\mu_t = \eta + (X_t - \eta)e^{-\kappa\Delta t}$$
$$\nu_t^2 = \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa\Delta t}\right)$$

The **log-likelihood function** of the observations is:

$$\mathcal{L}(\kappa, \eta, \sigma) = -\frac{1}{2}\sum_{i=1}^{N}\left[\log(2\pi\nu_t^2) + \frac{(x_i - \mu_{i-1})^2}{\nu_t^2}\right]$$

Maximizing $\mathcal{L}$ numerically yields the MLE estimators $(\hat{\kappa}, \hat{\eta}, \hat{\sigma})$.

Analytical MLE formulas. Following Appendix B of Baviera and Santagostino (2019), under regularly spaced observations with time step Δt, the MLE estimators for the OU parameters (κ, η, σ) admit the following closed-form expressions.
Define:

- $X_0 := \{x_0, \ldots, x_{N-1}\}$ and $X_1 := \{x_1, \ldots, x_N\}$
- $m_0 := \frac{1}{N}\sum_{i=0}^{N-1} x_i, \quad m_1 := \frac{1}{N}\sum_{i=1}^{N} x_i$
- $S_{00} := \sum_{i=0}^{N-1}(x_i - m_0)^2, \quad S_{11} := \sum_{i=1}^{N}(x_i - m_1)^2, \quad S_{01} := \sum_{i=1}^{N}(x_i - m_1)(x_{i-1} - m_0)$

Then:

$$\hat{\phi} := \frac{S_{01}}{S_{00}}, \quad \hat{\kappa} := -\frac{1}{\Delta t}\log(\hat{\phi})$$

$$\hat{\eta} := \frac{m_1 - \hat{\phi}m_0}{1 - \hat{\phi}}$$

$$\hat{\sigma}^2 := \frac{2\hat{\kappa}}{1 - \hat{\phi}^2} \cdot \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \hat{\eta} - \hat{\phi}(x_{i-1} - \hat{\eta})\right)^2$$

Stationary variance and standard deviation. The stationary variance ζ and standard deviation Σ of the OU process are then given by:

$$\hat{\zeta} = \frac{\hat{\sigma}^2}{2\hat{\kappa}}, \quad \hat{\Sigma} = \sqrt{\hat{\zeta}} = \frac{\hat{\sigma}}{\sqrt{2\hat{\kappa}}}$$

These quantities play a key role in determining the scaling of the trading bands and the critical transaction cost $c^*(\ell)$ in subsequent analysis.

Bootstrap procedure. After obtaining the MLE estimates, we simulate $M = 10{,}000$ synthetic OU paths under the estimated parameters and re-estimate the parameters on each simulated path. This yields bootstrap distributions $\{\kappa^{(i)}\}_{i=1}^{M}$, $\{\eta^{(i)}\}$ and $\{\sigma^{(i)}\}$ from which we compute 95% confidence intervals using empirical quantiles.

Results – 9:00 to 16:00 NYT. After filtering the data within this window and removing 4 antipersistent outliers, we obtain the following:

MLE Estimates and 95% Bootstrap Confidence Intervals

Parameter	Estimate	95% CI Lower	95% CI Upper
κ	23.7370	14.3586	52.6155
η (%)	−0.9772	−1.94042	−0.01040
σ (%)	10.25	9.9887	10.5196

Results – 8:00 to 16:00 NYT. With the earlier time window, the same 4 IS outliers are detected and excluded. The results are:

MLE Estimates and 95% Bootstrap Confidence Intervals

Parameter	Estimate	95% CI Lower	95% CI Upper
κ	24.0466	14.9976	53.2414
η (%)	−0.9804	−1.9688	−0.0297
σ (%)	10.35	10.1059	10.5941

2.4. Interpretation

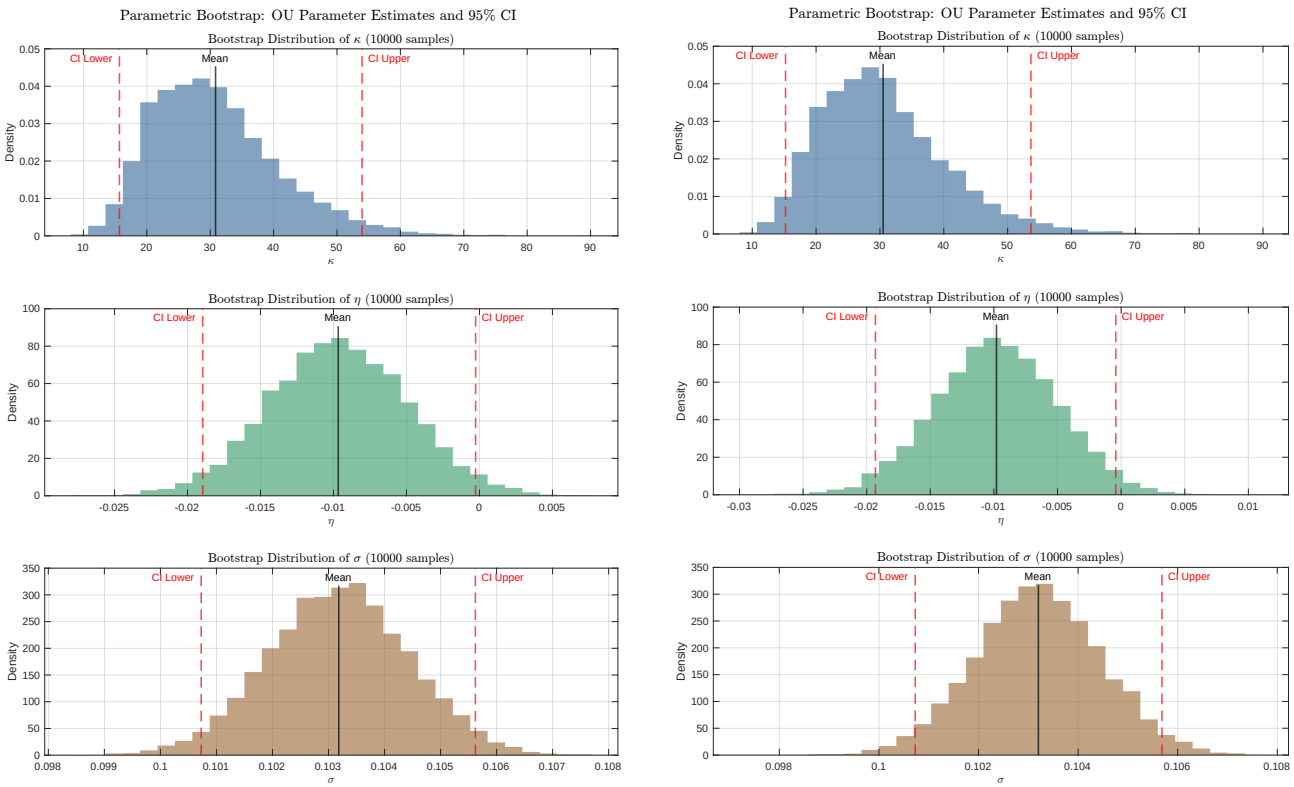
The estimates for both time windows are extremely consistent, showing that:

- The OU model is **robust** to the intraday time filter choice;
- The process is strongly mean-reverting ($\hat{\kappa} \approx 25$), The characteristic decay time, or *half-life*, is defined as the time required for a deviation from the mean to decay by 50%. It is given by:

$$t_{1/2} = \frac{\ln(2)}{\kappa} \approx \frac{0.693}{25} \approx 0.02\,\text{minutes} \approx 40\,\text{mins}.$$

  This indicates extremely mean reversion.
- The diffusion parameter σ is tightly bounded around 10.3%;
- The stationary mean η is negative and significantly different from zero.

Graphical summary. Figure 2 shows the bootstrap distributions of the estimated OU parameters κ, η, and σ for the two time windows considered. Each panel displays the empirical distribution from $M = 10{,}000$ bootstrap replications. Vertical dashed red lines denote the 2.5% and 97.5% percentiles, while the solid black line represents the sample mean.



(a) Bootstrap distributions for 9:00–16:00 NYT

(b) Bootstrap distributions for 8:00–16:00 NYT

Figure 2: Parametric bootstrap results for OU parameters. Each column shows the distributions for κ, η, and σ under the two intraday windows.

3. Question c

Objective. In this section we estimate the expected long-run return of the trading strategy with stop-loss $l = -1.645$ under three leverage configurations: $f = 1$, $f = 5$, and $f = f^*$ (optimal). We use the OU parameters calibrated In-Sample in Question b, and evaluate the Out-of-Sample (OS) returns over the last 3 months of the dataset, excluding the daily 17:00–20:00 NYT interval and the outliers.

Transaction costs. The total log transaction cost C_t at time t is estimated from the bid-ask spreads as:

$$C_t = \log\left(\frac{\text{Ask}_{\text{HO}}}{\text{Bid}_{\text{HO}}}\right) + \log\left(\frac{\text{Ask}_{\text{LGO}}}{\text{Bid}_{\text{LGO}}}\right)$$

The average cost over the IS period is:

$$C := \mathbb{E}[C_t] = \texttt{mean}(C_t) \approx 9.24\% \cdot \Sigma$$

where $\Sigma = \frac{\hat{\sigma}}{\sqrt{2\hat{\kappa}}}$ is the stationary standard deviation of the OU process.

Stop-loss and optimal thresholds. We fix the stop-loss at $L = -1.645 \cdot \Sigma$ and optimize the trading bands d, u by numerically maximizing the long-run return $\mu(d, u; f)$ using the formulas from Baviera and Santagostino (2019). The optimal thresholds depend on both f and c.
The optimization is done for:

- Constant leverage $f = 1$

3.1. Out-of-Sample Annualized Returns

To evaluate the real-world performance of our trading strategy, we compute the realized long-run return over the Out-of-Sample (OS) period using the average log-return:

$$\mu_t = \frac{1}{t} \log\left(\frac{W_t}{W_0}\right),$$

where W_t is the portfolio wealth at the end of the OS window.
In our setup, W_t updates only when the spread X_t moves from the entry level D to either the take-profit U or stop-loss L, with:

$$W_{t_{i+1}} = W_{t_i}(1 + f v_i), \quad v_i = \begin{cases} e^{U-D-C} - 1, & \text{if exited at } U, \\ e^{L-D-C} - 1, & \text{if exited at } L, \end{cases}$$

where f is the leverage and C the estimated transaction cost.
Let N^+ and N^- be the number of trades exiting at U and L, respectively. Then:

$$W_t = W_0(1 + f v^+)^{N^+}(1 + f v^-)^{N^-}, \quad \mu_t = \frac{1}{t}\left[N^+ \log(1 + f v^+) + N^- \log(1 + f v^-)\right].$$

This can be extended to the short side by symmetrizing the thresholds and including the corresponding return terms. We apply this computation to the OS dataset using the optimal thresholds (d^*, u^*), stop-loss ℓ, and transaction cost c estimated from IS data. The resulting annualized realized return is reported below, alongside the optimized thresholds and expected long-run return for **leverage l=1**.

Quantity	**Estimate**	**95% CI**
d^* $[\sigma]$	-0.922	$[-0.932, -0.913]$
u^* $[\sigma]$	0.597	$[0.569, 0.650]$
Expected μ^* (annual)	10.4%	[8.64%, 13.48%]
Out-of-Sample return	20.9%	–

3.2. Graphical illustration

Figures 3 and 4 provide a visual summary of the optimized trading strategies.
Figure 3 shows the log-ratio series R_t along with the corresponding entry, exit, and stop-loss levels for both the LONG and SHORT strategies under different leverage configurations: $f = 1$, $f = 5$, and $f = f^*$. All bands are computed in stationary units (Σ) and transformed back to real price levels.

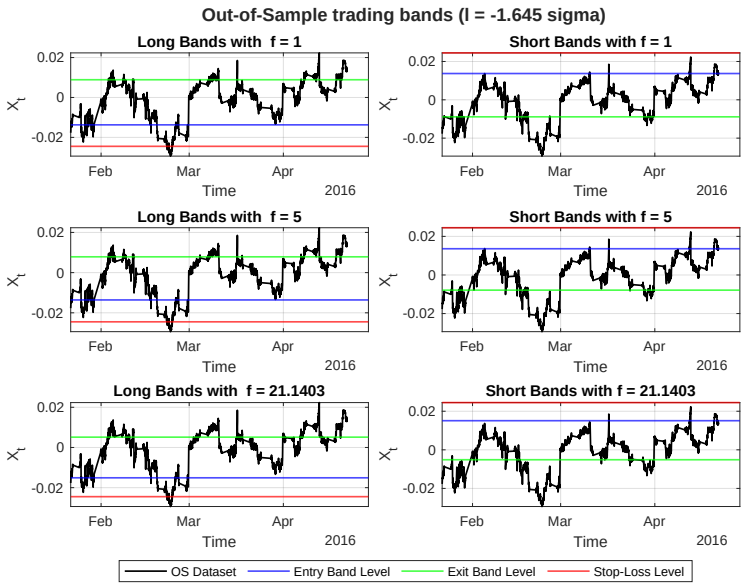


Figure 3: Price path R_t with trading bands in σ units.

To highlight the sensitivity of the optimal bands to transaction costs, Figure 4 plots the values of $|d|$ and u as a function of the cost c (in Σ units), for fixed stop-loss level $l = -1.645$. The green vertical line indicates the estimated cost C/Σ from our data. The red and blue markers show the values used in our backtest.

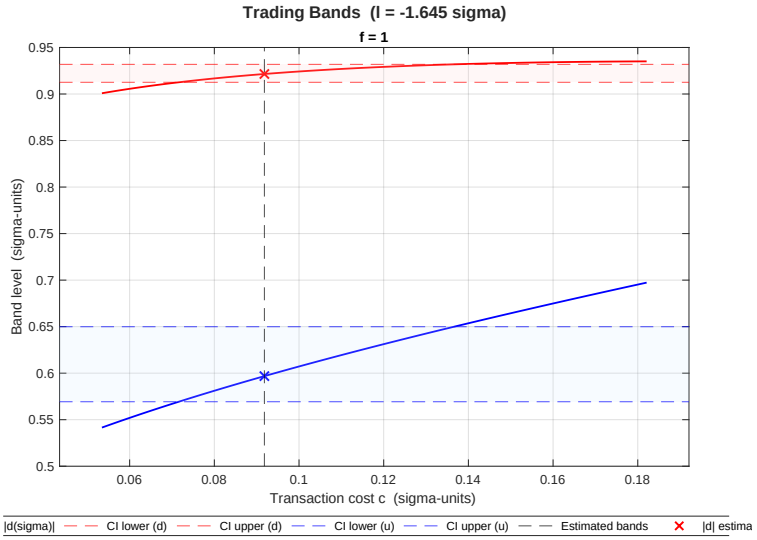


Figure 4: Optimal trading bands $|d|$ (entry) and u (exit) as a function of transaction cost C/Σ under stop-loss $l = -1.645$.

4. Question d

Objective. In this section we evaluate the impact of leverage on the optimal trading bands and the long-run return μ under a fixed stop-loss $l = -1.645$. Following the approach in Baviera and Santagostino (2019), we compute the thresholds d^* and u^* and the corresponding return for different values of leverage $f \in \{1, 2, 5, f^*\}$, using the **closed-form expression**.

Model parameters. We use the stationary volatility $\Sigma = \hat{\sigma}/\sqrt{2\hat{\kappa}}$ and mean-reversion timescale $\theta = 1/\hat{\kappa}$ calibrated in Question b. The transaction cost c is expressed in Σ units.

Results. With the exact closed-form expression, the results are:

Leverage	d^* $[\sigma]$	u^* $[\sigma]$	**Expected** μ	**OS Return**
$f = 1$	-0.922	0.597	10.4%	20.9%
CI	$[-0.932,\ -0.913]$	$[0.569,\ 0.649]$	$[8.62\%,\ 13.46\%]$	–
$f = 2$	-0.918	0.578	20.2%	57.0%
CI	$[-0.930,\ -0.907]$	$[0.544,\ 0.639]$	$[16.62\%,\ 26.26\%]$	–
$f = 5$	-0.913	0.528	45.9%	131.0%
CI	$[-0.928,\ -0.901]$	$[0.479,\ 0.607]$	$[37.01\%,\ 60.84\%]$	–
$f^* = 21.14$	-1.014	0.345	**104.3**%	**306.9**%
CI	$[-1.029,\ -0.989]$	$[0.294,\ 0.447]$	$[73.78\%,\ 162.13\%]$	$f^* \in [17.27,\ 25.14]$

4.1. Graphical illustration

Figure 5 shows the long-run return μ as a function of leverage f, along with the optimal point f^*. The function is concave as expected, with diminishing marginal gains and a unique maximum.

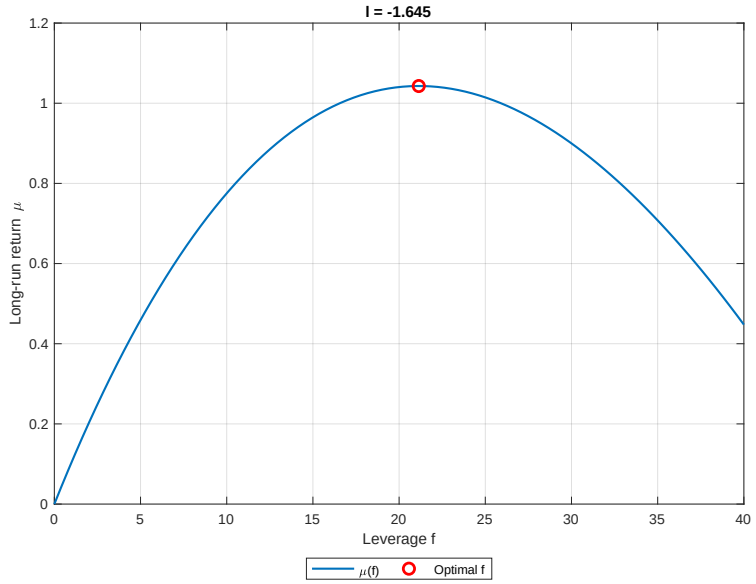


Figure 5: Long-run return μ as a function of leverage f, computed with the closed-form expression (flag = 2). Red dot indicates the optimal f^*.

5. Question e

Objective. In this section we investigate the impact of different stop-loss levels on the trading strategy's performance. Specifically, we fix the transaction cost and consider three stop-loss levels corresponding to quantiles of the stationary distribution:

$$l \in \{-1.282,\ -1.960,\ -2.326\} \quad (10\%,\ 2.5\%,\ \text{and}\ 1\%\ \text{quantiles})$$

For each stop-loss level, we compute the Out-of-Sample return under leverage $f = 1$, $f = 5$, and the **optimal** leverage f^*.

Results summary. We reported below the Out-of-Sample return just for f^* for each level of stop-loss (full result in the `MATLAB`):

Stop-loss l	f^*	**95% CI** f^*	**Expected** μ^*	**95% CI** μ^*	**OS Return**
-1.282	9.654	[7.296, 9.731]	24.4%	[15.67%, 25.31%]	137.5%
-1.960	28.225	[23.150, 37.487]	188.2%	[124.48%, 345.49%]	334.7%
-2.326	30.862	[24.466, 43.191]	260.7%	[168.02%, 507.51%]	554.8%

As expected, moderate levels of stop-loss (e.g., $l = -1.96$) balance the trade-off between limiting losses and capturing mean-reverting profits. Too loose a stop ($l = -2.326$) reduces the frequency of profitable trades, while too tight a stop ($l = -1.282$) increases the number of early exits and may result in negative return under optimal leverage.

Optimal leverage and trading bands. We also report the optimal leverage f^* and corresponding trading thresholds (in Σ units) for each l:

Stop-loss l	d^* $[\sigma]$**[95% CI]**	u^* $[\sigma]$**[95% CI]**
-1.282	-0.764 $[-0.789,\ -0.721]$	0.428 [0.364, 0.553]
-1.960	-1.106 $[-1.116,\ -1.098]$	0.299 [0.251, 0.393]
-2.326	-1.084 $[-1.125,\ -1.057]$	0.266 [0.216, 0.359]

5.1. Comment

The table reports the optimal leverage f^* and associated entry (d^*) and exit (u^*) thresholds, along with their 95% confidence intervals, for three different stop-loss levels l. As the stop-loss becomes more conservative (i.e., more negative), the optimal leverage increases, allowing the strategy to exploit wider excursions of the mean-reverting process.

Consistently, the entry threshold $|d^*|$ increases, signaling that the strategy waits for deeper deviations before entering, while the exit threshold u^* decreases slightly, enabling quicker profit-taking. The confidence intervals for f^* widen as l decreases, reflecting increased uncertainty due to the more aggressive positioning required. Notably, the bands remain well-separated for all cases, indicating stable and robust strategy calibration.

5.2. Reflections and Limitations

Although the theoretical model proves robust in controlled conditions, applying it to real-world high-frequency data requires caution. Some assumptions—such as the stationarity of the OU process or the constancy of transaction costs—may not hold under market stress. Moreover, the stop-loss parameter, while effective, remains sensitive to volatility spikes. Future work could investigate adaptive thresholding or regime-switching dynamics.

6. Efficient Confidence Interval Estimation via Interpolation and Discontinuity Handling

6.1. Bootstrap and the Motivation for Interpolation

To estimate confidence intervals (CIs) for trading bands and performance metrics such as the long-run return μ and the optimal leverage f^*, we apply a parametric bootstrap over the OU parameters (k_i, σ_i, η_i). A basic approach would require re-optimizing the trading bands (d^*, u^*) and f^* for each of the M bootstrap replications, resulting in substantial computational cost.
However, the return per unit of calendar time admits the decomposition:

$$\mu_i = \frac{1}{\theta_i} \cdot \mu_{\mathrm{OU}}(\Sigma_i), \quad \text{with} \quad \theta_i = \frac{1}{k_i}, \quad \Sigma_i = \frac{\sigma_i}{\sqrt{2k_i}}.$$

This structure implies that optimization depends only on Σ_i, not on θ_i, allowing for a grid-based precomputation over Σ. The resulting estimates are then interpolated across bootstrap samples.
Given a bootstrap draw i, we evaluate:

$$d_i = \mathrm{Interp}(\Sigma_i), \quad u_i = \mathrm{Interp}(\Sigma_i), \quad \mu_i = \frac{1}{\theta_i} \cdot \mathrm{Interp}(\mu_{\mathrm{OU}}(\Sigma_i)),$$

and, in the case of optimized leverage,

$$f_i^* = \mathrm{Interp}(f^*(\Sigma_i)).$$

This strategy provides a major speed-up while still capturing the bootstrap distribution of the estimators.

6.2. Accuracy and Validity of Interpolation

Interpolation is justified by the empirical smoothness of the mappings $\Sigma \mapsto d^*, u^*, \mu^*, f^*$. When leverage f is fixed, these functions are smooth and well-defined over the full parameter domain. In contrast, when f is optimized, discontinuities can arise where $f^* = 0$, corresponding to unprofitable strategies where trading is not undertaken.
In smooth regimes, standard linear interpolation yields second-order accuracy:

$$|g(\Sigma) - \tilde{g}(\Sigma)| \leq \frac{1}{8}\Delta\Sigma^2 \cdot \max_{\Sigma} |g''(\Sigma)|.$$

To balance accuracy and efficiency, we fix the grid spacing to $\Delta\Sigma = 10^{-4}$. Given the relatively narrow range of Σ values encountered in practice, this choice ensures numerical stability while remaining substantially faster than full re-optimization on each sample.

6.3. Discontinuities in f^*, Their Origin, and Statistical Treatment

For optimized leverage, the function $\Sigma \mapsto f^*$ may exhibit sharp discontinuities, especially for shallow stop-loss levels. In such cases, the expected return becomes non-positive, leading the optimizer to select $f^* = 0$, effectively disabling the trading strategy.
This creates jumps in the mappings:

$$\Sigma \mapsto f^*(\Sigma), \quad \mu^*(\Sigma), \quad d^*(\Sigma), \quad u^*(\Sigma),$$

which violate the smoothness assumptions of interpolation. To mitigate this, we explicitly identify the location of the profitability threshold in the Σ-grid. For bootstrap samples near or within the discontinuity, we bypass interpolation and compute the relevant quantities through direct optimization. In the remaining (profitable) region, interpolation remains accurate.
Confidence interval estimation proceeds as follows:

- If $f_i^* = 0$, we assign: $\mu_i = 0$, $d_i^* = \text{NaN}$, $u_i^* = \text{NaN}$,
- Confidence intervals for μ and f^* include all samples (including those with $f_i^* = 0$),
- Confidence intervals for d^* and u^* are computed only on samples with $f_i^* > 0$,
- The share of profitable samples is reported as:

$$\text{Profitable Share} = \frac{1}{M} \sum_{i=1}^{M} \mathbf{1}\left\{f_i^* > 0\right\}.$$

This approach ensures robustness of the statistical inference while preserving the interpretability of the estimated bands.

6.4. Visual Evidence: Shallow vs. Deep Stop-Loss

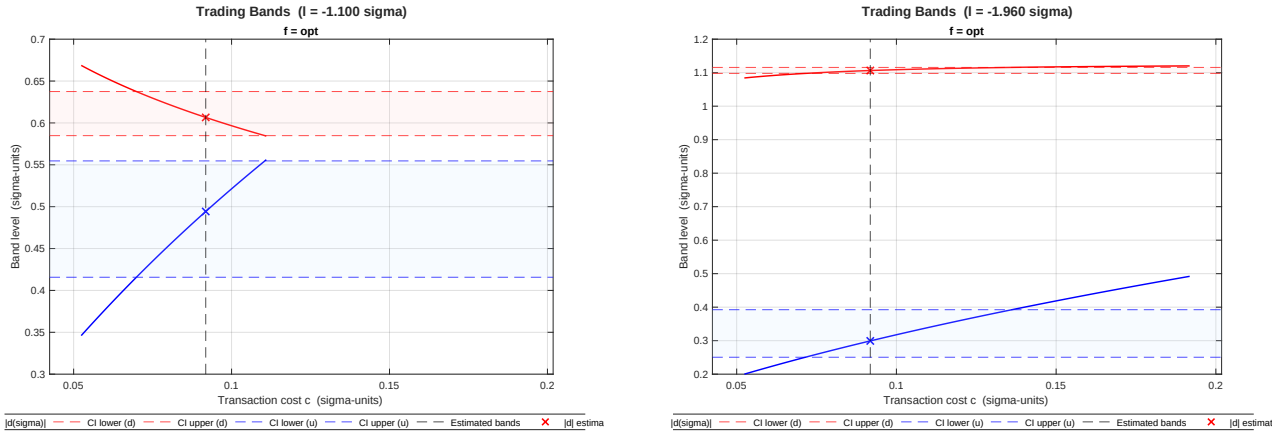
We illustrate the effect of stop-loss depth on the continuity of estimated quantities. Consider:

- A shallow stop-loss $\ell = -1.100$, where a non-negligible fraction of bootstrap samples yield $f^* = 0$, resulting in discontinuities;
- A deep stop-loss $\ell = -1.960$, for which all samples remain profitable and the mappings are smooth.

Comparison of optimal trading band estimates under shallow vs. deep stop-loss levels

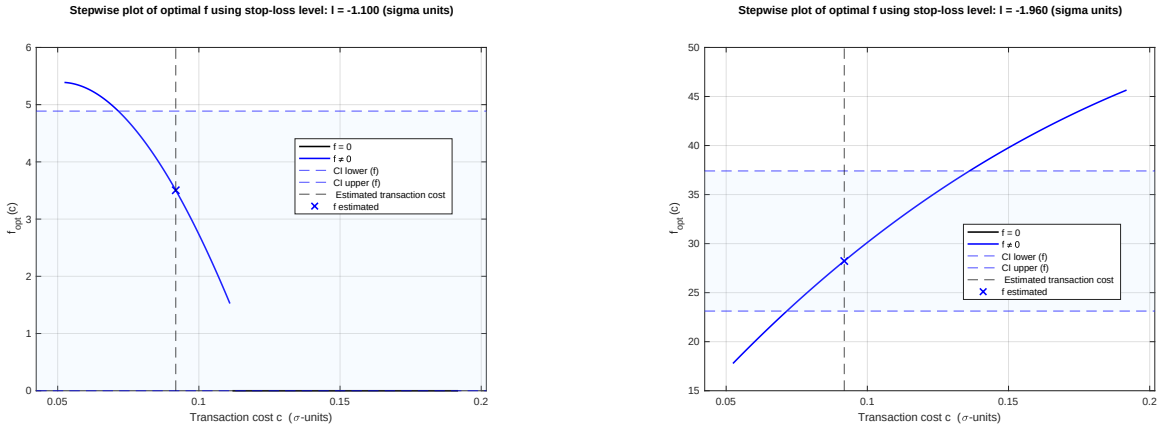
Metric	$\ell = -1.100$ (shallow)	$\ell = -1.960$ (deep)
u^*	0.494	0.299
d^*	-0.607	-1.106
μ^*	0.034	1.882
95% CI for u^*	[0.416, 0.555]	[0.252, 0.393]
95% CI for d^*	[-0.637, -0.585]	[-1.116, -1.098]
95% CI for μ^*	[0.00%, 6.17%]	[125.87%, 344.70%]
95% CI for f^*	[0, 4.89]	[23.30, 37.46]
Profitable share	74.740%	100%

Hereafter we show plots of optimal trading bands and of f^* with respect to transaction cost c (Σ units) in the range of $c = C/\Sigma$ with $\Sigma \in (\Sigma_{min}, \Sigma_{max})$ from M statistical bootstrap parameters.



(a) $\ell = -1.200$ (b) $\ell = -1.960$

Figure 6: Trading bands $d^*(c)$, $u^*(c)$ under different stop-loss levels. In the shallow case bands are not defined for every c.



(a) $\ell = -1.200$ (b) $\ell = -1.960$

Figure 7: Optimal leverage $f^*(c)$ under different stop-loss levels. Stepwise discontinuities and flat zones appear at shallow ℓ, while the deep case remains smooth.

The stop-loss levels shown in Figures 6 and 7, particularly the shallow value $\ell = -1.100$, were intentionally chosen despite being unrealistic from a practical trading perspective. Our goal was not to propose viable trading rules but to stress-test the robustness of our algorithm under edge cases where structural discontinuities in the optimization problem become visible. By contrasting this with the smoother case at $\ell = -1.960$, we clearly highlight the type of instability that can arise and demonstrate how our interpolation and confidence interval methodology handles such issues rigorously and consistently.

6.5. Summary

This framework provides a computationally efficient yet statistically robust method for estimating confidence intervals under parameter uncertainty. By interpolating over smooth regions and selectively re-optimizing near discontinuities, we maintain both speed and reliability. The distinction between profitable and non-profitable samples is explicitly handled, ensuring meaningful inference without masking structural breaks.

7. Risk Evaluation via Bootstrapped Out-of-Sample Return (Conceptual Framework)

The following outlines a concise conceptual extension of the existing workflow to include a bootstrapped out-of-sample (OS) risk evaluation. Though not implemented in this study, this approach would enable a more comprehensive assessment of the strategy's robustness under model uncertainty, trading band variation, and leverage.

7.1. Bootstrapped Parameter Estimation from IS Data

Using a bootstrap procedure on in-sample (IS) data, we generated M resampled paths to capture estimation uncertainty. For each path, we derived the corresponding trading bands and leverage:

$$\{(\Sigma_i, \theta_i, d_i, u_i, f_i, \mu_{\text{IS},i})\}_{i=1}^{M}$$

Each configuration represents a variant of the strategy, reflecting parameter uncertainty.

7.2. Pathwise Simulation on the Fixed OS Dataset

Unlike previous OS evaluations that relied on a single estimated strategy and summary statistics, this conceptual extension simulates the full evolution of portfolio wealth, taking into account the sequencing of trades, using the M bootstrapped configurations. Each strategy is applied to the same OS dataset with the following trading logic:

- Enter a trade when the process breaches the threshold $d_i \cdot \Sigma_i$.
- Exit at $u_i \cdot \Sigma_i$ (take-profit) or $l \cdot \Sigma_i$ (stop-loss).
- Returns are applied multiplicatively to capital, incorporating leverage f_i and transaction costs.
- Simulation ends if wealth drops below a predefined threshold (e.g., 50% of initial capital).

This process produces a distribution of wealth trajectories, capturing variability due to parameter uncertainty.

7.3. Risk Metrics

From the simulated wealth paths, the following risk metrics can be computed:

- **Annualized OS return:**
$$\mu_{OS_i} = \frac{\log(W_i^{\text{final}})}{T_{\text{OS}}}$$
- **Risk of ruin:** fraction of simulations ending below the wealth threshold.
- **Return confidence interval:** empirical quantiles (e.g., 2.5%, 97.5%).
- **Maximum drawdown:** worst peak-to-trough decline per wealth path.

7.4. Brief Interpretation

- A narrow CI on μ_i means the strategy is relatively insensitive to parameter noise; a wide CI signals fragility.
- A high ruin rate suggests that some reasonable parameter draws can decimate the account, even if average return is positive.
- Large drawdowns in the tail imply a real danger of severe capital dips—prompting consideration of smaller position sizes or tighter stops.

8. Question g

Application to Governative Bond Futures

Objective. In this final section, we assess the robustness and generalizability of the statistical arbitrage framework by applying it to a new asset class: **governative bond futures**. Unlike the energy futures previously analyzed, these instruments exhibit distinct characteristics in terms of liquidity, volatility, and potential cointegration structure. This makes them a compelling testbed for evaluating the out-of-sample reliability of the proposed strategy.

Methodology. We consider the following three pairs of governative bond futures:

`IKA -- OATA` `OEA -- OATA` `IKA -- RXA`

For each pair, we construct the log-price ratio time series using mid-quote data and apply the same modeling and calibration procedure introduced earlier. The dataset is partitioned into two regimes:

- **In-Sample (IS):** four-month interval for model estimation.
- **Out-of-Sample (OS):** two-month interval for strategy evaluation.

The spread dynamics are modeled as a discretized Ornstein–Uhlenbeck (OU) process. Parameters are estimated via **Maximum Likelihood Estimation (MLE)** with **bootstrapped confidence intervals** using $M = 10{,}000$ samples and significance level $\alpha = 5\%$.

Trading strategies are implemented under a fixed stop-loss level of $l = -1.645$ (in units of spread standard deviation) and the corresponding **optimal leverage** f^*, determined by maximizing the expected long-run return. We report both the theoretical performance metrics and the realized Out-of-Sample returns.

Remark. The results reported refer to the case $l = -1.645$ and the corresponding optimal leverage f^*. For results under different stop-loss thresholds and leverage levels, please refer to the accompanying `MATLAB` file.

8.1. Pair IKA vs OATA

Descriptive Statistics and Outlier Detection

The IKA-OATA pair was analyzed over an in-sample (IS) period of 45,444 observations and an out-of-sample (OS) period of 21,099 observations. No outliers were detected in the IS data, while 1,103 were identified in the OS sample (5.23%), suggesting a shift in regime. The IQR-based filter may misclassify legitimate market movements as noise in the OS period, especially if volatility has changed. Notably, a clear structural break is visually evident just after March 2025—the IS/OS boundary—as shown in Figure 8, where the log price ratio between the two assets drops sharply. This regime change raises concerns about the robustness of the in-sample calibration when applied to the OS period.

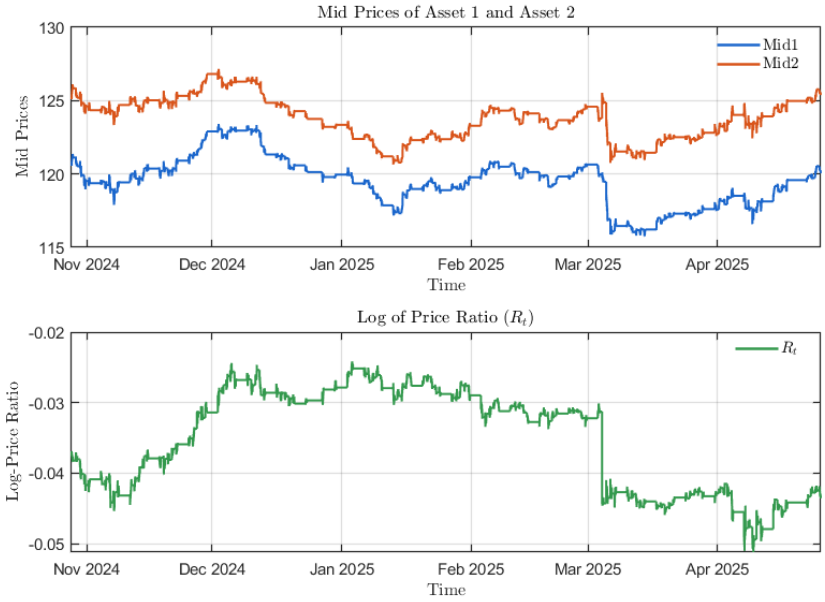


Figure 8: IKA vs OATA prices and the log of price ratio

OU Model Estimation

The log-ratio dynamics were modeled via a discretized Ornstein-Uhlenbeck (OU) process. Maximum likelihood estimates and bootstrapped confidence intervals are reported below:

MLE Estimates and 95% Bootstrap Confidence Intervals

Parameter	Estimate	95% CI Lower	95% CI Upper
κ	24.2064	13.1221	78.8314
η (%)	−3.1115	−3.6074	−2.6349
σ (%)	3.4300	3.4029	3.4476

The estimation is stable, with only one bootstrap sample removed due to a negative κ, corresponding to 0.01%.

Optimal Band Strategy and Leverage

Using a stop-loss of $l = -1.645$, the optimal thresholds and leverage were computed by maximizing the risk-adjusted return. Results are reported in Table 1.

Table 1: Optimal Trading Bands and Risk-Adjusted Return

Metric	Estimate	95% CI Lower	95% CI Upper
Upper band u^*	0.169	0.1412	0.2508
Lower band d^*	−1.075	−1.0907	−1.0427
Risk-adjusted return μ^*	1.560	0.9026	4.1804
Optimal leverage f^*	85.0430	65.9188	132.7205

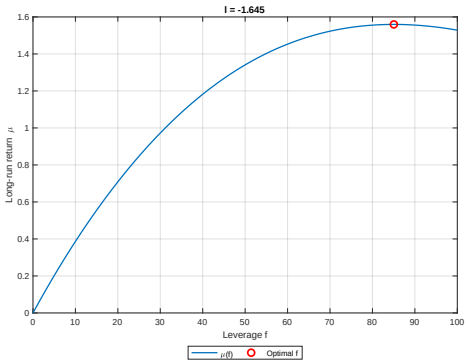


Figure 9: Expected return $\mu(f)$ as a function of leverage f for IKA-OATA.

Graphical Analysis

Figure 10 displays the log-price ratio R_t over the out-of-sample period, with superimposed entry/exit thresholds and stop-loss. Despite the calculated bands, no trades were triggered during the OS period.

Figure 10: Out-of-sample log-price ratio R_t and optimal trading bands for IKA-OATA.

Out-of-Sample Performance

The strategy yields a negative out-of-sample return of **-1.901**, despite active trading and strong in-sample results. This contradicts earlier interpretations and confirms a clear **structural break** between the IS and OS periods, as visually evident in Figure 10. Model parameters estimated in-sample fail to generalize, leading to systematic losses.

8.2. Pair OEA vs OATA

Descriptive Statistics and Outlier Detection

The OEA-OATA pair was analyzed over an in-sample (IS) period of 45,444 observations and an out-of-sample (OS) period of 21,099 observations. No outliers were detected in either dataset, suggesting a high degree of stationarity and regime consistency between the IS and OS periods.

OU Model Estimation

The spread dynamics were modeled via a discretized Ornstein-Uhlenbeck (OU) process. Maximum likelihood estimates and bootstrapped confidence intervals are reported below:

MLE Estimates and 95% Bootstrap Confidence Intervals

Parameter	Estimate	95% CI Lower	95% CI Upper
κ	31.9097	17.8721	90.3723
η (%)	−5.0942	−5.5494	−4.6561
σ (%)	4.2100	4.1797	4.2346

Optimal Band Strategy and Leverage

Using a stop-loss of $l = -1.645$, the optimal thresholds and leverage were computed by maximizing the risk-adjusted return. Results are presented in Table 2.

Table 2: Optimal Trading Bands and Risk-Adjusted Return

Metric	Estimate	95% CI Lower	95% CI Upper
Upper band u^*	0.163	0.1382	0.2299
Lower band d^*	−1.078	−1.0926	−1.0501
Risk-adjusted return μ^*	2.085	1.2382	5.0497
Optimal leverage f^*	80.3634	63.0138	120.2348

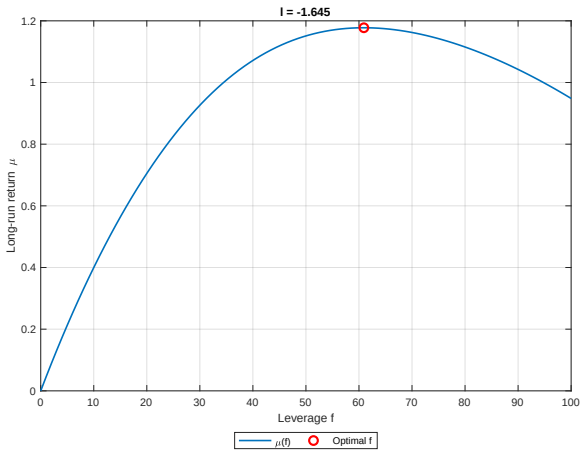


Figure 11: Expected return $\mu(f)$ as a function of leverage f for OEA-OATA.

Graphical Analysis

Figure 12 displays the log-price ratio R_t over the out-of-sample period, with superimposed entry/exit thresholds and stop-loss. The strategy was triggered, but resulted in a negative out-of-sample return.

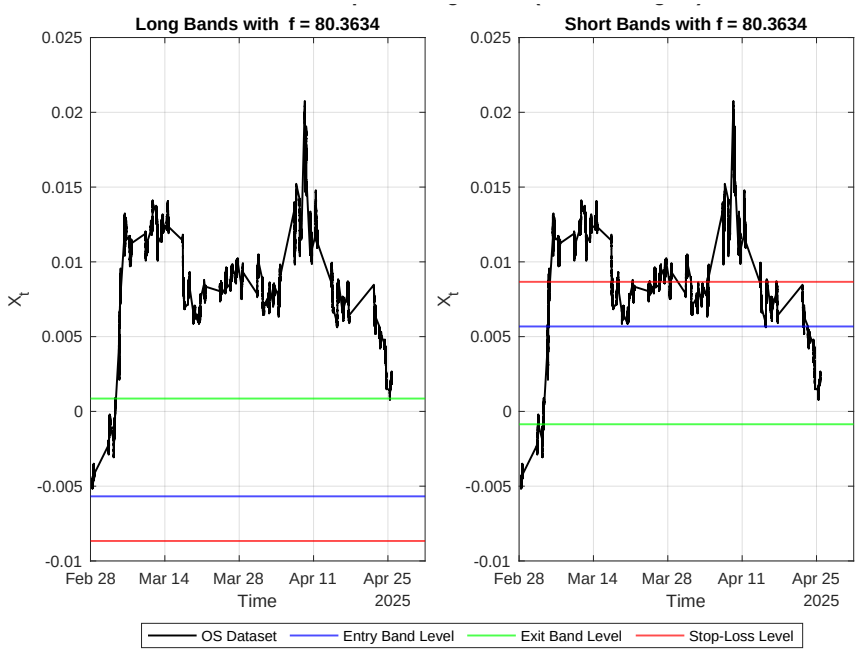


Figure 12: Out-of-sample log-price ratio R_t and optimal trading bands for OEA-OATA.

Out-of-Sample Performance

The strategy yielded an out-of-sample return of **-3.814**, despite strong in-sample calibration and confidence intervals. This discrepancy may be attributed to:

- Short-lived excursions beyond the band levels leading to adverse timing;
- Increased market noise or shifts in microstructure not reflected in the IS;
- Sub-optimal reversion strength despite a high κ.

8.3. Pair IKA vs RXA

Descriptive Statistics and Outlier Detection

The IKA-RXA pair was analyzed over an in-sample (IS) period of 45,444 observations and an out-of-sample (OS) period of 21,099 observations. No outliers were detected in either dataset, suggesting regime consistency and stable statistical properties between the two periods.

OU Model Estimation

The spread dynamics were modeled via a discretized Ornstein-Uhlenbeck (OU) process. Maximum likelihood estimates and bootstrapped confidence intervals are reported below:

MLE Estimates and 95% Bootstrap Confidence Intervals

Parameter	Estimate	95% CI Lower	95% CI Upper
κ	75.3574	49.6214	143.7934
η (%)	−10.0205	−10.2268	−9.8177
σ (%)	4.5000	4.4719	4.5307

Optimal Band Strategy and Leverage

Using a stop-loss of $l = -1.960$, the optimal thresholds and leverage were computed by maximizing the risk-adjusted return. Results are reported in Table 3.

Table 3: Optimal Trading Bands and Risk-Adjusted Return

Metric	Estimate	95% CI Lower	95% CI Upper
Upper band u^*	0.200	0.1748	0.2499
Lower band d^*	−1.061	−1.0727	−1.0430
Risk-adjusted return μ^*	4.515	3.1524	7.6429
Optimal leverage f^*	107.9866	91.7222	136.7187

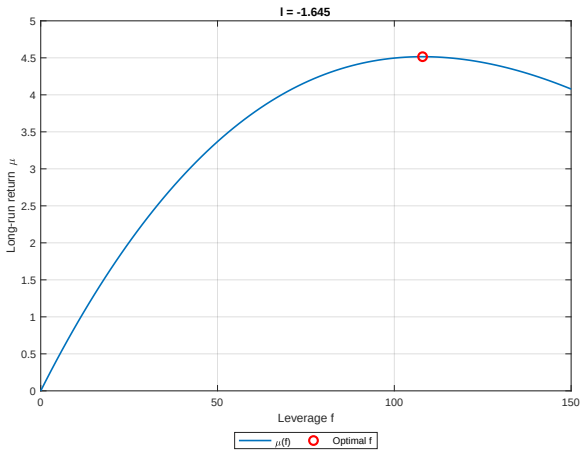


Figure 13: Expected return $\mu(f)$ as a function of leverage f for OEA-OATA.

Graphical Analysis

Figure 14 displays the log-price ratio R_t over the out-of-sample period, with superimposed entry/exit thresholds and stop-loss. The strategy was active and generated a profit.

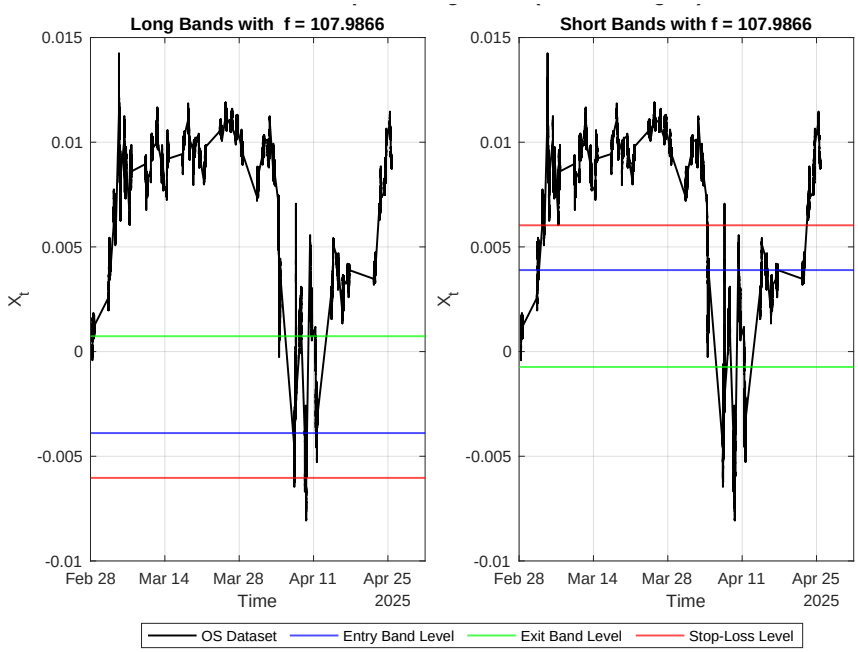


Figure 14: Out-of-sample log-price ratio R_t and optimal trading bands for IKA-RXA.

Out-of-Sample Performance

The strategy yielded an out-of-sample return of **2.439**, indicating successful execution within the defined trading bands. The strong in-sample calibration and absence of detected outliers support the strategy's robustness in this pair.

9. Question h

9.1. Cryptocurrency Data Extraction Process

To conduct high-frequency pairs trading analysis in the cryptocurrency market, we designed and implemented a parallelized data extraction pipeline based on the Binance API. The objective was to retrieve **1-minute OHLCV data** (Open, High, Low, Close, Volume) over a 6-month historical period for the **ten most liquid cryptocurrency pairs against USDT**.
The selected symbols, ranked by liquidity, are listed below:

BTCUSDT ETHUSDT BNBUSDT SOLUSDT XRPUSDT
ADAUSDT AVAXUSDT DOGEUSDT LINKUSDT MATICUSDT

For consistency and liquidity purposes, we restrict our analysis to **BTC-X_Crypto** combinations, using BTCUSDT as the base asset in all pairs.

The analysis has been conducted in Python in file: `Final_Project_Crypto_Group_8B.ipynb`

9.2. Statistical Testing for Ornstein-Uhlenbeck Assumption

To evaluate the mean-reversion potential of each price spread, we apply the **Augmented Dickey-Fuller (ADF) test** (see Said and Dickey (1984)) to the log price ratio between each BTC-based pair and its counterpart. The ADF test assesses the null hypothesis of a unit root in the time series, which would imply non-stationarity. Rejection of the null hypothesis (i.e., a p-value below a chosen significance threshold, such as 0.05) supports the assumption of stationarity.
Stationarity is a fundamental requirement for modeling a spread as an **Ornstein-Uhlenbeck (OU) process**, which underpins many pairs trading strategies. If the spread is stationary, it indicates that price deviations from the historical mean are likely to be temporary and mean-reverting—making the pair suitable for statistical arbitrage.

Example: BTCUSDT-XRPUSDT

The results of the ADF test applied to the log spread between BTCUSDT and XRPUSDT are reported below:

ADF Statistic	−3.3054
p-value	0.0146

ADF Test Results for BTCUSDT-XRPUSDT Spread

Since the p-value is below the 0.05 threshold, we **reject the null hypothesis** of a unit root. This supports the stationarity of the spread and validates the use of an OU process to model its dynamics.

9.3. Model calibration – parameter estimates

Maximum-likelihood estimation of the Ornstein–Uhlenbeck parameters on 1-minute data yields the results in Table 4. Confidence intervals are computed via parametric bootstrap with 10,000 replications.

Table 4: Estimated OU parameters (time unit: minutes)

Parameter	Estimate	95% Confidence Interval
Mean-reversion speed κ (min^{-1})	78.612	[55.246, 131.232]
Long-run mean η	10.539	[10.496, 10.585]
Volatility σ	1.054	[1.051, 1.057]

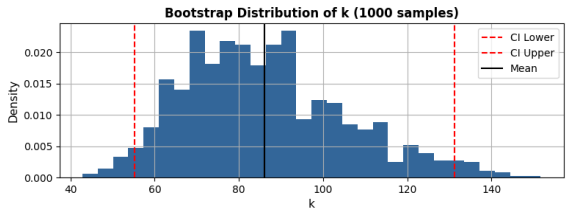
Interpretation.

- **Speed of mean reversion** The parameter κ governs how fast the process X_t reverts to its long-run mean. Since it multiplies a time derivative, its unit is 1/time.
  The characteristic decay time, or *half-life*, is defined as the time required for a deviation from the mean to decay by 50%. It is given by:

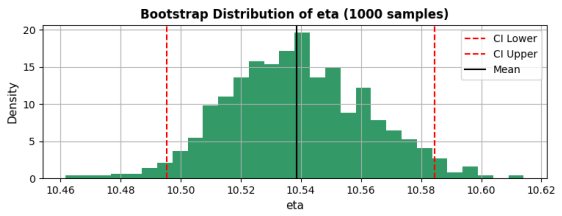
$$t_{1/2} = \frac{\ln(2)}{\kappa} \approx \frac{0.693}{78.61} \approx 0.0088\,\text{minutes} \approx 13\,\text{mins}.$$

  This indicates extremely fast mean reversion.
- **Long-run mean** The parameter $\eta = 10.539$ is the equilibrium level to which Y_t reverts. If the process moves away from this value, the pullback force is proportional to κ.
- **Volatility** The diffusion coefficient $\sigma = 1.054$ determines the magnitude of random fluctuations. Even with fast mean reversion, a high σ leads to frequent excursions away from the mean. The stationary variance of the process is $\sigma^2/(2\kappa)$.

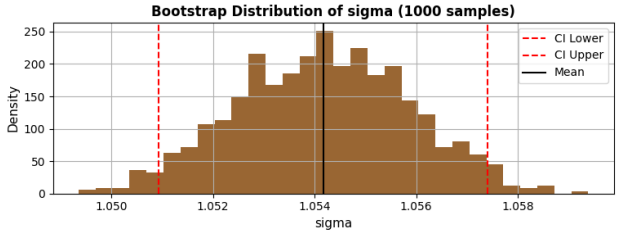
The statistical bootstrap leads to the following results :



(a) κ confidence interval and distribution



(b) η confidence interval and distribution



(c) σ confidence interval and distribution

Figure 15: Distribution and CI intervals of the estimated parameters

9.4. Analysis

SL	Lev	d^*	d_{lo}	d_{hi}	u^*	u_{lo}	u_{hi}	μ	μ_{lo}	μ_{hi}	f^*	f_{lo}	f_{hi}
-1.282	1	0.980	0.965	0.990	0.514	0.508	0.519	1.819	1.533	2.329			
-1.282	2	0.962	0.948	0.974	0.399	0.387	0.415	3.065	2.506	4.072			
-1.282	5	1.039	0.987	1.079	0.177	0.146	0.218	4.477	3.291	6.713			
-1.282	opt	1.044	1.024	1.057	0.168	0.151	0.183	4.480	3.323	6.954	5.155	4.444	6.383
-1.645	1	0.828	0.827	0.832	0.434	0.432	0.438	2.500	2.101	3.220			
-1.645	2	0.782	0.780	0.790	0.301	0.279	0.332	4.391	3.600	5.819			
-1.645	5	0.964	0.864	1.067	0.126	0.126	0.144	7.127	5.321	10.512			
-1.645	opt	1.117	1.112	1.121	0.133	0.125	0.142	7.303	5.331	11.620	6.258	5.312	7.925
-1.960	1	0.630	0.625	0.642	0.335	0.330	0.347	3.119	2.620	4.020			
-1.960	2	0.563	0.553	0.582	0.195	0.167	0.232	5.655	4.659	7.449			
-1.960	5	0.767	0.646	0.907	0.002	0.002	0.031	9.738	7.243	14.289			
-1.960	opt	0.964	0.950	0.981	0.029	0.025	0.032	9.980	7.250	15.993	6.156	5.204	7.859
-2.326	1	0.438	0.430	0.451	0.251	0.243	0.263	3.822	3.210	4.925			
-2.326	2	0.368	0.354	0.392	0.125	0.096	0.165	7.150	5.926	9.353			
-2.326	5	0.601	0.457	0.796	-0.081	-0.082	-0.023	13.099	9.640	19.123			
-2.326	opt	0.779	0.762	0.803	-0.062	-0.062	-0.058	13.312	9.644	21.411	5.768	4.882	7.368

SL = -1.282

- Optimal trading bands: entry at $d^* \approx 1.044$, exit at $u^* \approx 0.168$
- Long-run return: $\mu \approx 4.48$ (95% CI: [3.32, 6.95])
- Optimal cost leverage: $f^* \approx 5.16$
- **Interpretation:** Mild stop-loss yields moderate returns and high tolerance for transaction costs.

SL = -1.645

- Bands widen: entry $d^* \approx 1.117$, exit $u^* \approx 0.133$
- Long-run return: $\mu \approx 7.30$ (95% CI: [5.33, 11.62])
- Optimal cost leverage: $f^* \approx 6.26$
- **Interpretation:** Tighter stop-loss improves risk-adjusted return, possibly by avoiding deep drawdowns.

SL = -1.960

- Entry threshold slightly lower: $d^* \approx 0.964$, exit narrows further: $u^* \approx 0.029$
- Long-run return: $\mu \approx 9.98$ (95% CI: [7.25, 15.99])
- Optimal cost leverage: $f^* \approx 6.16$
- **Interpretation:** Strategy becomes more conservative but still achieves significant expected returns.

SL = -2.326

- Entry threshold drops further: $d^* \approx 0.779$, exit approaches zero: $u^* \approx -0.062$
- Long-run return: $\mu \approx 13.31$ (95% CI: [9.64, 21.41])
- Optimal cost leverage: $f^* \approx 5.77$
- **Interpretation:** Deep stop-loss yields the best return, but requires highly efficient execution due to narrow exit bands.

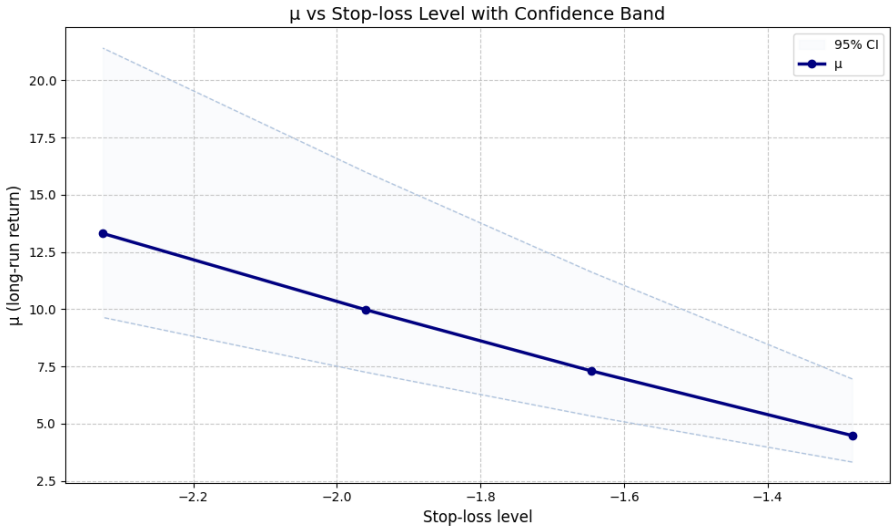


Figure 16: μ (long-run return) versus the stop-loss level (SL), with 95% confidence intervals.

Stricter stop-loss levels (more negative thresholds) consistently lead to higher long-run returns (μ). The optimal return, $\mu \approx 13.31$, occurs around -2.33σ, compared to $\mu \approx 4.48$ at -1.28σ. However, confidence intervals widen with tighter stops, reflecting increased sensitivity to estimation and execution errors. This highlights a trade-off between return maximization and robustness: deeper stops improve performance but demand greater precision and lower transaction costs. These findings underscore the importance of disciplined risk control and accurate band calibration in mean-reverting strategies.

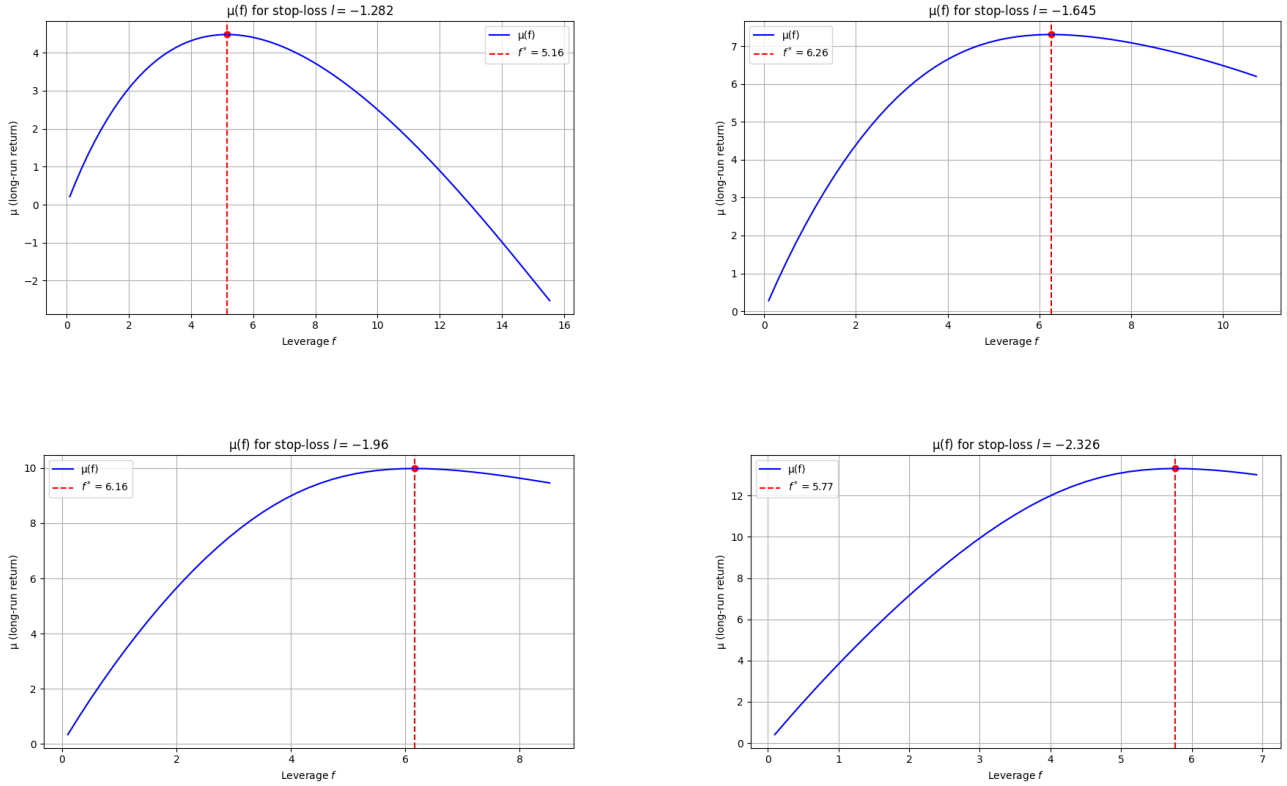


Figure 17: μ as a function of leverage f for different Stop-loss levels

When plotting the long-run return μ as a function of leverage, we observe a concave relationship. Initially, as leverage increases, the expected return also rises, reflecting the amplified gains from exploiting the mean-reverting process. However, beyond a certain point, the return starts to decline due to increased exposure to adverse price movements, transaction costs, and tighter exit conditions. This behavior highlights the existence of an **optimal leverage level**—a peak in the μ vs. leverage curve—beyond which the compounding negative effects outweigh the benefits of additional leverage. Practically, this implies that while leverage can enhance performance, excessive leverage introduces fragility and reduces the long-run growth rate, underscoring the importance of risk-aware position sizing.

10. General Conclusions and Critical Assessment

In this project, we have replicated, extended, and applied the Ornstein–Uhlenbeck (OU) stop–loss arbitrage framework introduced by Baviera and Santagostino (2019) to a variety of markets, such as energy commodities, government-bond futures, and cryptocurrency pairs. Our empirical results largely confirm the theoretical predictions, but they also reveal several important practical considerations. Below, we discuss the main strengths and weaknesses of the method.

10.1. Strengths

- **Solid theoretical basis.** The OU model with an explicit stop–loss has the advantage of yielding closed-form expressions for optimal entry and exit thresholds, as well as for expected profit. This analytical tractability allows for fast, transparent optimization without resorting to extensive numerical simulations.
- **Rigorous estimation of parameters.** We have employed maximum-likelihood estimation for the OU parameters (κ, σ, η), and supplemented these estimates with 10,000 bootstrap replications to obtain tight confidence intervals. This approach proved reliable across diverse datasets, from 30-minute energy spreads to 1-minute cryptocurrency quotes.
- **Explicit consideration of transaction costs.** By embedding a critical cost parameter directly into the optimization problem, the framework automatically adjusts the trading bands to discourage over-trading in markets with wide bid–ask spreads or high fees. This feature helps protect the strategy's edge in environments where costs can quickly erode profit.
- **Transparent risk control.** The stop–loss level ℓ, measured in units of σ, provides a mechanical way to limit tail risk. We have quantified how varying ℓ affects both the expected return μ and the Kelly-type optimal leverage $f^\star$. This clarity makes it straightforward to calibrate the approach according to an investor's risk tolerance.
- **Applicability across multiple asset classes.** Through case studies on heating-oil versus gas-oil spreads, Treasury-futures pairs, and BTC-based crypto pairs, we have shown that the pipeline is largely asset-agnostic. Once the spread is defined and data are available, the same OU stop–loss machinery can be applied with minimal modifications.

10.2. Weaknesses

- **Dependence on stationarity.** The core OU assumption—that the spread is mean-reverting—can break down during regime shifts, such as monetary-policy announcements, exchange halts, or crypto hard forks. Without regular stationarity checks, one risks deploying the model on spreads that are drifting rather than reverting, which can lead to persistent losses.
- **Vulnerability to microstructure noise.** At very high frequencies (one minute or below), estimates of the mean-reversion speed κ often become inflated due to noise. This leads to unrealistically tight entry/exit bands and can generate an excessive number of trades, eroding profitability once real-world latencies and costs are accounted for.
- **Underestimation of execution costs.** In our backtests, we included a reasonable "critical cost" parameter, but real-world slippage, tiered commission schedules, and maker-taker rebates can be more complicated than a single cost estimate. In thinly traded crypto pairs or bond futures with wide bid–ask spreads, execution costs may exceed the assumed threshold, reducing or even eliminating the strategy's edge.
- **Occasionally unrealistic leverage.** The optimal leverage $f^\star$ can exceed $20\times$ in energy and bond futures or suggest sub-second half-lives in crypto. In practice, margin requirements, exchange limits, and latency constraints make such aggressive leverages impractical. Traders must cap leverage to feasible levels, at the expense of some theoretical return.
- **Limited protection against extreme events.** A stand-alone stop-loss does not guard against overnight price gaps, exchange closures, or sudden flash crashes. In live trading, it is often necessary to augment the OU stop–loss with additional safeguards—such as circuit breakers, dynamic hedges, or options—to cover those "tail" scenarios.

- **Scalability and crowding risks.** As capital flows into stat-arb strategies like this one, the available arbitrage opportunities tend to shrink. Market participants who attempt to scale the strategy must be aware that the statistical edge will decay as the pool of similar traders grows, particularly in less liquid markets.

10.3. Final Remarks

Overall, the OU stop–loss approach represents an elegant, reproducible, and pedagogically clear framework for statistical arbitrage on price spreads. Our empirical tests confirm that the theoretical predictions hold up under realistic conditions, but the transition from backtest to live trading invariably brings operational challenges. In particular, the following safeguards are recommended for any real-world implementation:

1. **Weekly rolling stationarity checks.** Implement regular Augmented Dickey–Fuller (ADF) tests to verify that the spread remains mean-reverting. Recalibrate OU parameters after each check to capture any evolving dynamics.
2. **Inclusion of actual trading costs.** Replace the "critical cost" placeholder with real commission schedules, anticipated slippage estimates, and funding or borrowing rates. Run out-of-sample simulations that incorporate these detailed costs to assess the net profitability more accurately.
3. **Stress-testing leverage.** Simulate scenarios with sudden volatility spikes or liquidity crunches to see how high-leverage positions behave under extreme conditions. Set conservative caps on $f^\star$ to prevent margin calls or forced liquidations during adverse market moves.

When these measures are in place—weekly recalibration to ensure stationarity, rigorous cost accounting, and prudent stress-testing—the OU stop–loss technique can serve as a robust component of a diversified stat-arb portfolio. It is capable of generating decorrelated alpha, provided that market conditions continue to align with the model's core assumptions about mean reversion and cost structure.

References

Baviera, R. and L. Santagostino (2019). "Stop-loss and leverage in optimal statistical arbitrage". In: *Energy Economics* 84.

ICE Futures Europe (2024). *Low Sulphur Gasoil Futures Contract Rules (Section J1)*. Technical contract specifications. URL: https://www.theice.com/publicdocs/contractregs/15_SECTION_J1.pdf.

Said, S.E. and D.A. Dickey (1984). "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order". In: *Biometrika* 71.3, pp. 599–607. DOI: 10.1093/biomet/71.3.599. URL: https://larrylisblog.net/WebContents/Financial%20Models/ADFTest.pdf.