Module Title: High Frequency Trading

Module Code: 7FNCE025W

Course: MSc Fintech and Business Analytics, Semester 2, 2023/2024

ASSIGNMENT 2: Pairs Trading Strategy

Table of Contents

A.Data Acquisition

As per the requirement we are asked to create a pairs trading strategy using the high frequency minutely data of Bitcoin and Litecoin, high frequency historical data is crucial for understanding market dynamics and finding short-term arbitrage opportunities. This data provides in depth view granularly of price movements and aids in statistical and mathematical analysis.

'yfinance' library was used to fetch the historical data for 'BTC' and 'LTC' on a complete day of 11th April 2024, starting from '00:00' to '11:59', with a '1 minute' interval to ensure high granularity.

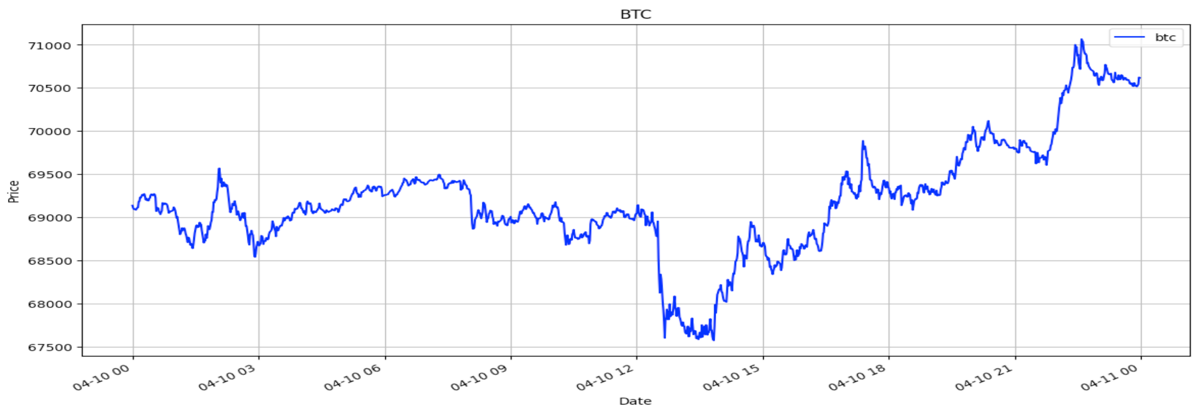

Figure 1: BTC Price

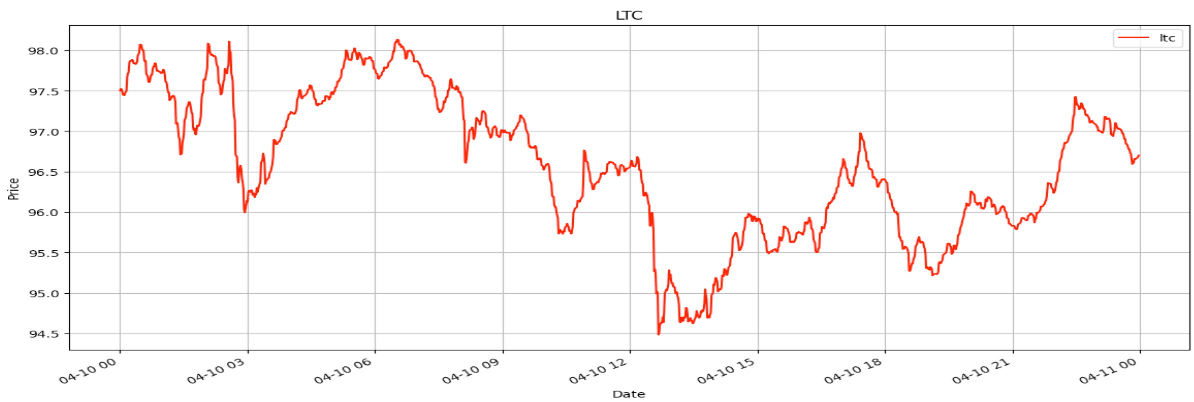

Figure 2: LTC Price

The data collected offers a detailed perspective on various metrics like opening price, closing price, highest, lowest and volume traded. Theses are essential for pairs trading strategy as it helps in exploring the relationship between assets with historical correlation. Potential trade opportunities might be found when the price ratio deviates from its historical average by analysing high-frequency data.

Pairs trading is a non-directional, relative value investing approach that looks for two linked financial products with comparable features, like stocks or funds. The idea is to find the cheap financial instrument and short the overvalued one. Furthermore, one of the most important things to note is that for pairs trading the assets need to be cointegrated. Instead of making money from the market's general direction, the objective is to profit from the convergence of the prices of the two instruments back to their historical relationship. (Göncü and Akyıldırım, 2016)

Since the p-value from the Engle-Granger cointegration test is grater than 0.05 we reject the null hypothesis and conclude that the given assets are not cointegrated but since we need to answer the following questions we might still go with the process. (Yip Chee-Yin and Hock Eam Lim, 2014)

```
from statsmodels.tsa.stattools import coint

# Test for cointegration
score, p_value, _ = coint(p1,p2)
print(f'Engle-Granger cointegration test p-value: {p_value}')
```

Engle-Granger cointegration test p-value: 0.9863752993443312

B. Estimating the Correlation Parameter (ρ)

The degree of correlation between two variables is indicated by the correlation coefficient, or ρ. Perfect positive correlation is represented by a value of +1, perfect negative correlation by a value of -1, and no correlation by a value of 0. It is an imperfect criterion for selecting pairs, as financial data is often not normally distributed, and correlation cannot fully capture the dependence structure. Cointegration analysis is sometimes used instead of or in addition to correlation to identify suitable pairs for trading. It calculates the degree and direction of a linear relationship between Litecoin (LTC) and Bitcoin (BTC) returns. Setting $\delta=1$, which relates to a particular component of the dynamics or risk preferences of our approach, is a fundamental assumption in our research. Assuming this, the emphasis switches to determining the correlation parameter ρ, which is important for modelling the assets' relationship and predicting future price fluctuations. (Bernstein and Nielsen, 2019)

The log returns of LTC and BTC are computed to get ρ. The required estimate is provided by the correlation between these log returns.

- Log returns for BTC and LTC

```
r1=np.log(btc['Adj Close']/btc['Adj Close'].shift(1))
r2=np.log(ltc['Adj Close']/ltc['Adj Close'].shift(1))
```

- Estimating ρ between BTC and LTC returns

```
rho=r1.corr(r2)
```

The computed ρ shows the minute-by-minute movement of the prices of BTC and LTC in relation to one another. Whereas a negative correlation points to inverse movements, a stronger positive correlation shows that the assets often move in the same direction. Stable and substantial correlations are important in pairs trading because methods frequently take advantage of the asset price ratio's reversion to its historical mean. Comprehending ρ enables one to ascertain the degree of correlation between BTC and LTC and to anticipate their future price ratio with confidence.

In summary, a critical step in the pair's trading method was done by assuming δ=1 and calculating ρ. Our statistical modelling relies heavily on the estimated correlation coefficient, which guides our trading decisions by pointing out price divergences and convergences between Bitcoin and Litecoin. Setting up a statistical framework to take advantage of chances when the price ratio of the asset considerably deviates from its predicted value.

C. Setting and Explaining a Reasonable Beta (β)

The beta (β) parameter is important when it comes to pairs trading since it gauges the relative volatility or systematic risk of an asset in comparison to the market or, in this example, the paired asset. It does this by analysing the asset's price movement in relation to another asset.

A decent estimation of beta is necessary for pairs trading strategies involving Bitcoin (BTC) and Litecoin (LTC) to comprehend how changes in BTC's price may impact changes in LTC's price, or vice versa. It is essential to comprehend this in order to forecast and profit from price changes.

A linear regression study that compares the returns of the two assets is commonly used to estimate beta. The scipy.stats module's linregress function is used in this method to determine beta (β), the slope of the regression line between the log prices of BTC and LTC. Here's the relevant code:

- linear regression between the log prices of BTC and LTC

```
result=linregress(np.log(btc['Adj Close']),np.log(ltc['Adj Close']))
```

- The slope of the regression line is our beta (β)

```
beta=result.slope
```

It is possible to ascertain how fast and in which direction LTC prices are anticipated to move in reaction to changes in BTC prices by running linear regression on the log prices of BTC and LTC. A greater beta would indicate that the price of LTC is more erratic in relation to changes in the price of BTC. On the other hand, a lower beta denotes reduced sensitivity. When developing strategies that may involve leveraging the more volatile asset or hedging against price fluctuations, the value of beta is helpful.

It's important to comprehend beta since it plays a key role in determining the hedging ratio—the amount of one asset that should be purchased or sold to protect against changes in the other's price. Regardless of the direction of the market, this ratio is essential for preserving a market-neutral stance and hoping to profit from the price ratio's convergence to its historical average.

We establish the groundwork for a more knowledgeable and successful pairs trading strategy by establishing a resonable beta based on past price movements of BTC and LTC. It makes it possible to comprehend and predict how much the price of LTC would react to a shift in the price of BTC, which improves decision-making throughout the strategy's development and implementation.

D. Discussing Strategy Inputs: S, μ, and σ

Several inputs are essential to the modelling and decision-making processes involved in the creation of any financial strategy, but pairs trading in particular. The most important ones are the asset's volatility (σ), expected returns (μ), and initial stock prices (S). These characteristics are the foundation of our mathematical and statistical modelling in the pairs trading strategy for Bitcoin (BTC) and Litecoin (LTC), affecting our trading decisions and forecasts.

1. **Initial Stock Prices (S)**:

The opening prices of BTC and LTC, respectively, on the first day of our data set are represented by **S1_0** and **S2_0**.

```
S1_0 = btc['Open'][0]
S2_0 = ltc['Open'][0]
```

2. **Expected Returns (μ)**:

The mean of each cryptocurrency's log returns is multiplied by the number of trading periods in our timeframe to determine the expected returns for BTC and LTC.

```
mu_1 = r1.mean()*1411
mu_2 = r2.mean()*1411
```

3. **Volatility (σ)**:

The standard deviation of each cryptocurrency's log returns, scaled by the square root of the number of trading periods, represents the annualised volatility for both Bitcoin and Litecoin.

```
sigma_1 = r1.std()*np.sqrt(1411)
sigma_2 = r2.std()*np.sqrt(1411)
```

The key components of our pairs trading method are the starting stock prices (S), expected returns (μ), and volatilities (σ) of Bitcoin and Ethereum. They are employed not only in the strategy's original design but also in its continuous implementation and risk control. Since these characteristics directly affect the strategy's predictions, possible rewards, and risks, their precise estimation and careful attention are essential to its success. These inputs will be examined frequently and modified as necessary to reflect shifting market circumstances and enhance our trading approach.

E. Finding Optimal Trading Weights

How much capital to put into each asset in a pair trading strategy is determined by optimal trading weights. These weights are essential for optimising returns while lowering risk in accordance with the objectives and limitations of the strategy. Considering the predicted returns, volatility, and correlation between the assets, the weights need to represent the relative attractiveness of each position. (Lintilhac and Tourin, 2016)

The following formulas determine the ideal weights (π_1 and π_2) for Bitcoin (S_1) and Litecoin (S_2) at every time interval, taking into account volatilities, predicted returns, and additional factors like as beta and correlation coefficient. Here's the procedure:

$$\pi_1^* S_1 = \frac{(\mu_1 + \delta z)}{\gamma(1-\rho^2)\sigma_1^2} + \frac{\delta(-2a(t)(\mu_1 + \delta z) - b(t))}{\gamma} - \frac{\rho\mu_2}{\gamma(1-\rho^2)\sigma_1\sigma_2}$$

$$\pi_2^* S_2 = \frac{\mu_2}{\gamma(1-\rho^2)\sigma_2^2} + \frac{\delta\beta(-2a(t)(\mu_1 + \delta z) - b(t))}{\gamma} - \frac{\rho(\mu_1 + \delta z)}{\gamma(1-\rho^2)\sigma_1\sigma_2}$$

$$\begin{aligned}
a(t) &= \frac{T-t}{2(1-\rho^2)\sigma_1^2}, \\
b(t) &= -\frac{1}{4}(\sigma_1^2 + \beta\sigma_2^2)\delta\frac{(T-t)^2}{(1-\rho^2)\sigma_1^2} - \frac{\rho\mu_2}{(1-\rho^2)\sigma_1\sigma_2}(T-t), \\
c(t) &= \frac{1}{2}\frac{\mu_2^2}{(1-\rho^2)\sigma_2^2}(T-t) + \frac{1}{4}(\sigma_1^2 + \beta\sigma_2^2 + 2\sigma_1\sigma_2\beta\rho)\delta^2\frac{(T-t)^2}{(1-\rho^2)\sigma_1^2} \\
&\quad + \frac{1}{4}\frac{\mu_2(\sigma_1^2 + \beta\sigma_2^2)\delta\rho}{(1-\rho^2)\sigma_1\sigma_2}(T-t)^2 + \frac{1}{24}(\sigma_1^2 + \beta\sigma_2^2)^2\delta^2\frac{(T-t)^3}{(1-\rho^2)\sigma_1^2}.
\end{aligned}$$

The ideal quantity of each asset to hold at any one time in order to maintain a position that is neutral to the market and optimise returns is represented by the trading weights. They are dynamic, reacting to shifts in the market's temperature, volatility, and other factors.

```
weightBTC = Pi_1[t]
weightLTC = Pi_2[t]
```

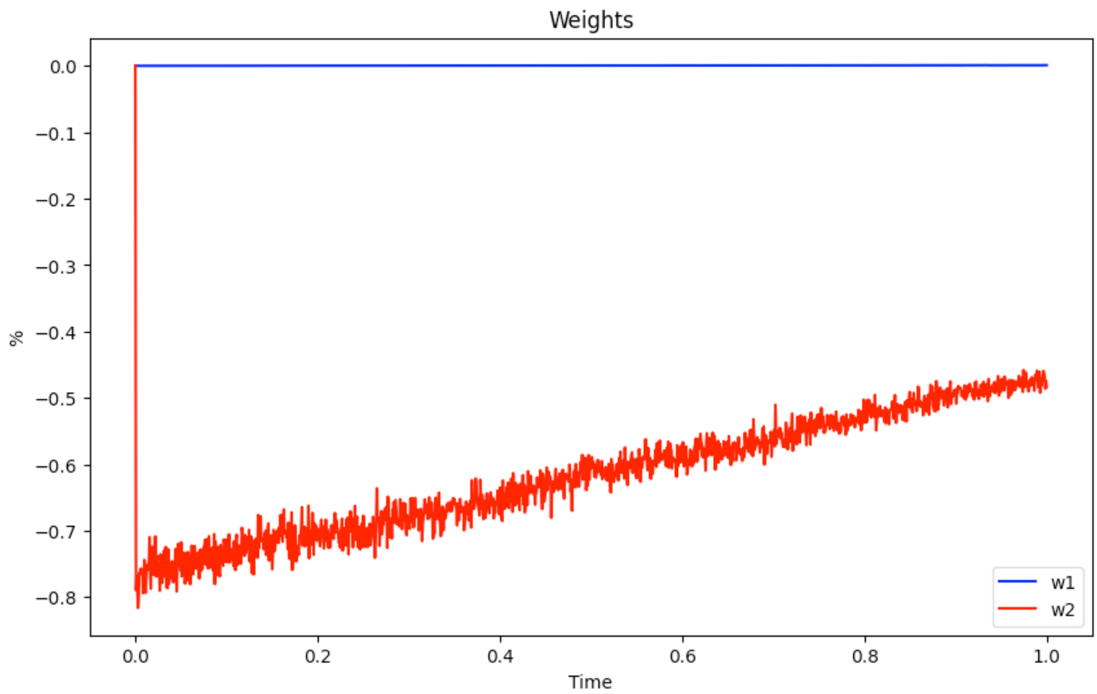

Figure 3: Dynamic Weights

Since the weights are dynamic, it is important to regularly check them and make adjustments in reaction to shifting market conditions. The way these weights are visualised makes it evident how the strategy distributes capital over time and adapts to the changing market conditions. This knowledge is essential for managing the pairs trading strategy both in the short and long term.

F. Finding the Optimal Cash Process

Given the ideal trading weights, the optimal cash process in pairs trading is the progressive distribution of financial resources among the assets. It's an essential component of the strategy, stating the amount of money required or invested in each position in order to keep the weights at the correct levels. This procedure aids in comprehending the capital needs and the positions' anticipated future monetary value.

```
cashBTC = Pi_1[t] * S1[t]
cashLTC = Pi_2[t] * S2[t]
```

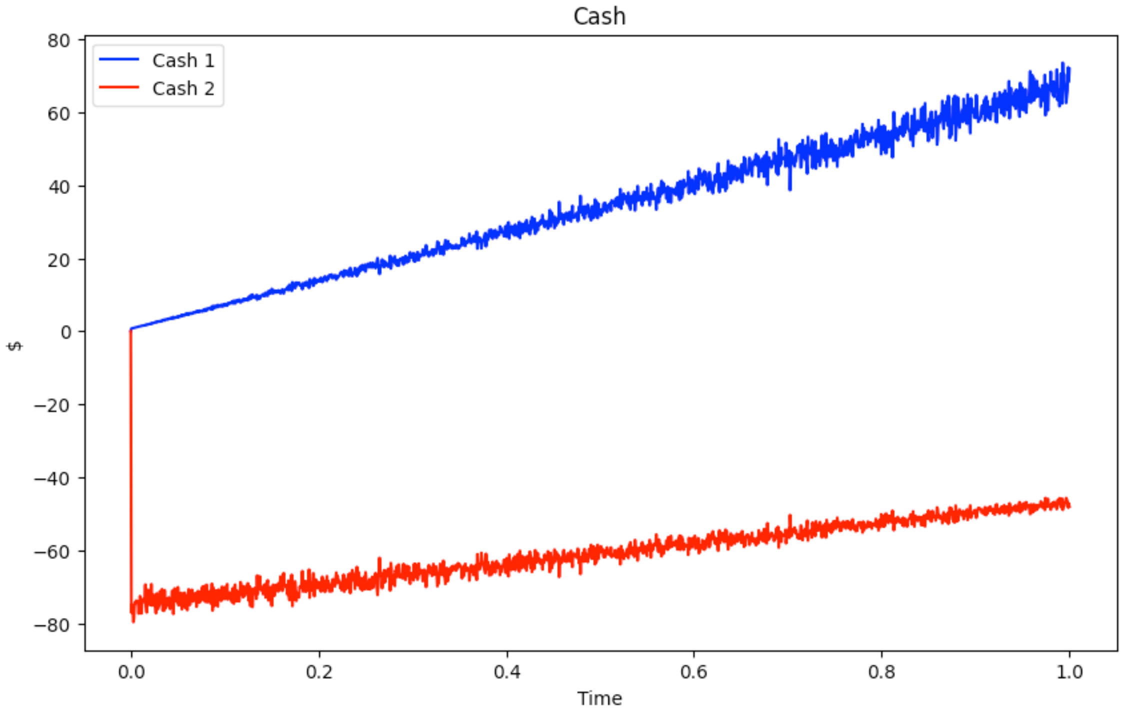

Figure 4: Dynamic Cash holding.

The actual dollar amount invested in each asset at any one time is reflected in the optimal cash process. Both the asset's price and the percentage of the portfolio devoted to it have a direct impact on it. Here are some conclusions and things to think about:

- **Capital Allocation**: It offers perceptions into the capital distribution across assets throughout time, assisting in the comprehension of exposure and possible capital needs.
- **Market Dynamics Response**: The cash process is dynamic and reacts to shifts in the weights and prices of the assets, showing how the strategy adapts to fresh market intelligence.
- **Strategy Execution**: The cash process tells the trader how much capital needs to be moved into or out of each position over time, therefore understanding it is essential to actually executing the strategy.

One of the most important steps in carrying out a pairs trading strategy is determining the best cash method. It offers a transparent picture of each position's cumulative monetary value, illustrating how the strategy distributes resources among the assets. Understanding the dynamics of capital allocation and controlling and modifying the strategy in response to market fluctuations are made easier with the aid of the cash process visualisation. Having this understanding is essential to keeping a profitable and balanced trading position when using the pairs trading method.

G. Comparing Results with Static Control Pair

To determine how effective a dynamic approach is in pairs trading, one must compare its performance against that of a static control pair. Instead of dynamically changing in reaction to market changes, a static control pair usually refers to a method where the weights (and hence the cash process) are maintained throughout the trading period. This comparison makes it easier to see how dynamic rebalancing adds value and determines whether better performance justifies higher transaction costs and complexity.

Setting Reasonable Static Control Factors:

It is crucial to establish appropriate factors or weights that are steady throughout time in order to preserve static control. These could be equal weighting, risk preferences, or historical averages. Assume for comparison's sake that we allocate 50% of our capital to Bitcoin and 50% to Litecoin over the course of the investment period, using equal weights for each asset as our static control.

- Static weights for comparison

```
staticWeightBTC = 0.5
staticWeightLTC = 0.5
```

- Static cash process

```
staticCashBTC = staticWeightBTC * S1
staticCashLTC = staticWeightLTC * S2
```

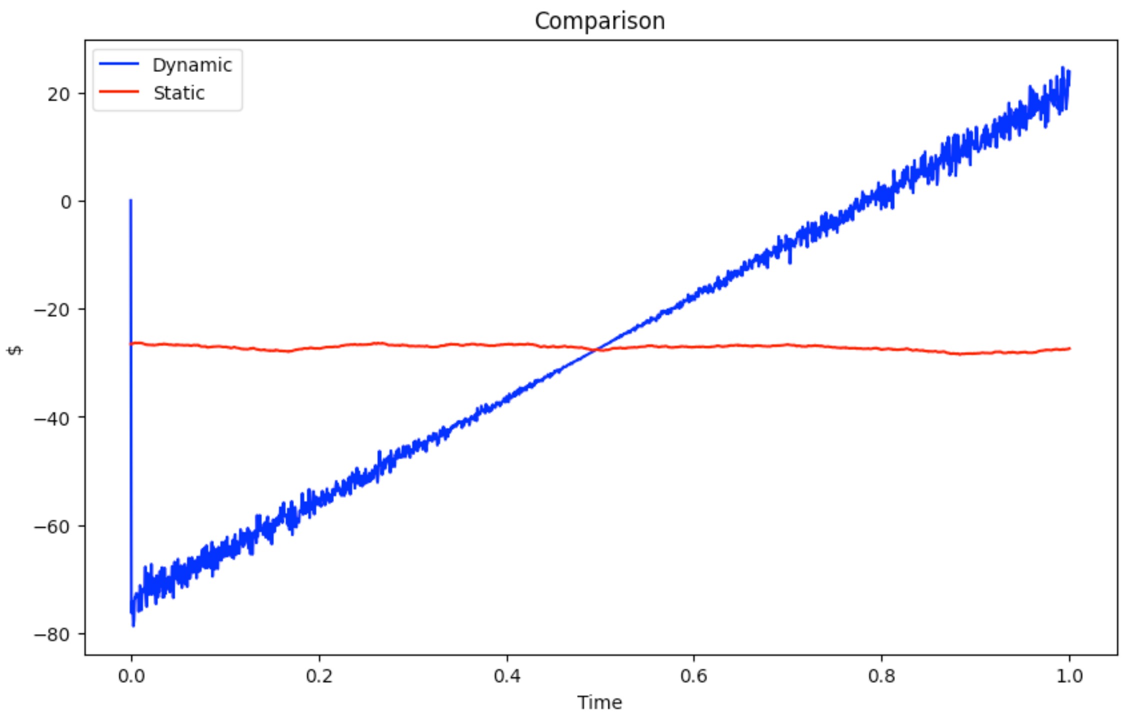

Figure 5: Returns Comparison

Evaluating the efficacy of the dynamic pairs trading technique requires comparing it to a static control pair. Traders can decide whether dynamic rebalancing is worthwhile and what the trade-offs are by establishing appropriate static control elements and tracking the overall amount of money invested over time. This comparison draws attention to the possible performance advantages as well as the costs, complexity, and risk management factors that are specific to dynamic versus static techniques. The comparison's findings will direct the pairs trading strategy's continued management and improvement.

H. Simulating Two Assets

With the help of price simulations, our pairs trading approach may be used in a safe, fictitious setting. This simulation might give you an idea of how the approach might work in different market scenarios. Following the asset price simulation, the correlation estimation, optimal trading weight computation, and optimal cash process determination will be repeated. Following this, the two optimal control factors will be shown and contrasted with a static control.

A stochastic process, such as Geometric Brownian Motion (GBM), which is a popular technique for modelling asset values in financial mathematics, will be used to simulate the price pathways of two assets.

The simulation enables us with:

• Testing the approach in multiple circumstances: Gaining insight into how the pairs trading strategy might operate in various market settings that are simulated using stochastic processes.

• Verifying the strategy's viability and profitability in a controlled setting before using it in actual trading is known as strategy validation.

• Visualisation of control factors: tracking the actions of the two simulated assets and evaluating the relative merits of a dynamic vs static strategy.

In this simulation we took the stock price s1 as 400 and s2 as 800, $\mu 1$=0.02, $\mu 2$=0.04, $\sigma 1$=0.1, $\sigma 2$=0.05, z_0=0.02, β=1, γ=,1 δ=0.5, ρ=-0.7.

Where,

$\mu 1$ and $\mu 2$ are the mean return of stock 1 and stock 2.

$\sigma 1$ and $\sigma 2$ are variance in returns of stock 1 and stock 2.

z_0 defines the stationarity of the pair at t=0

β defines the long-term correlation between the price of the stocks.

ρ defines the correlation between the return of the stocks.

γ defines the risk appetite.

δ defines the speed of mean reversion to the long-term average.

Here the if the value of β increases the stocks prices tend to converge whereas when it decreases the stock prices starts to diverge. Similarly, the higher the value of γ the less is the risk appetite and vice versa.

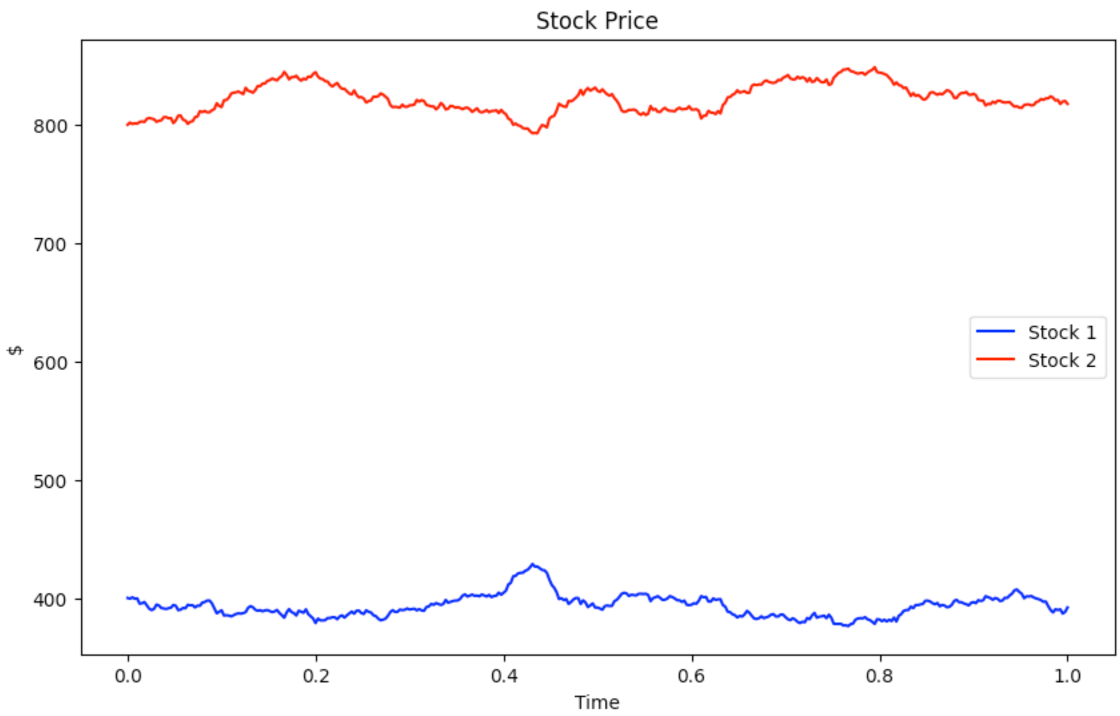

Figure 6: Stock Price.

The simulation strategy based on the inputs provided resulted in a better return for the dynamic for the dynamic strategy that involved **dynamic weights and dynamic cash as the two optimal control factors**, which tells us that the strategy works well with the dynamic weights and dynamic cash for optimized and better returns.

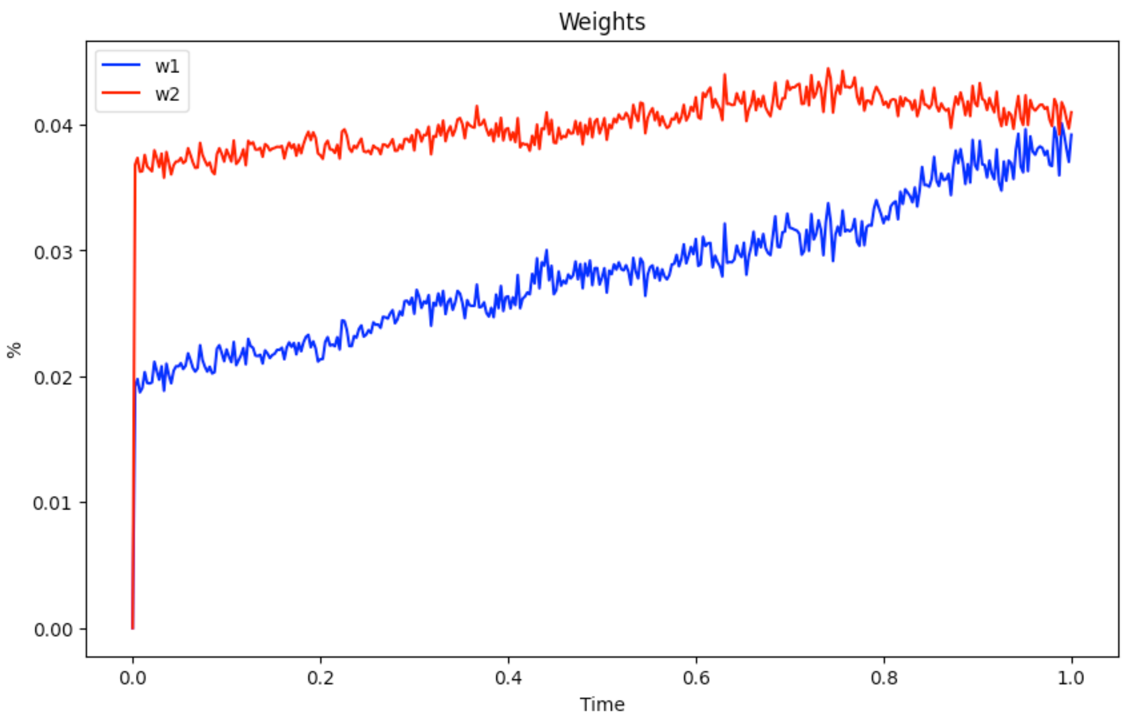

Figure 7: Dynamic Weights

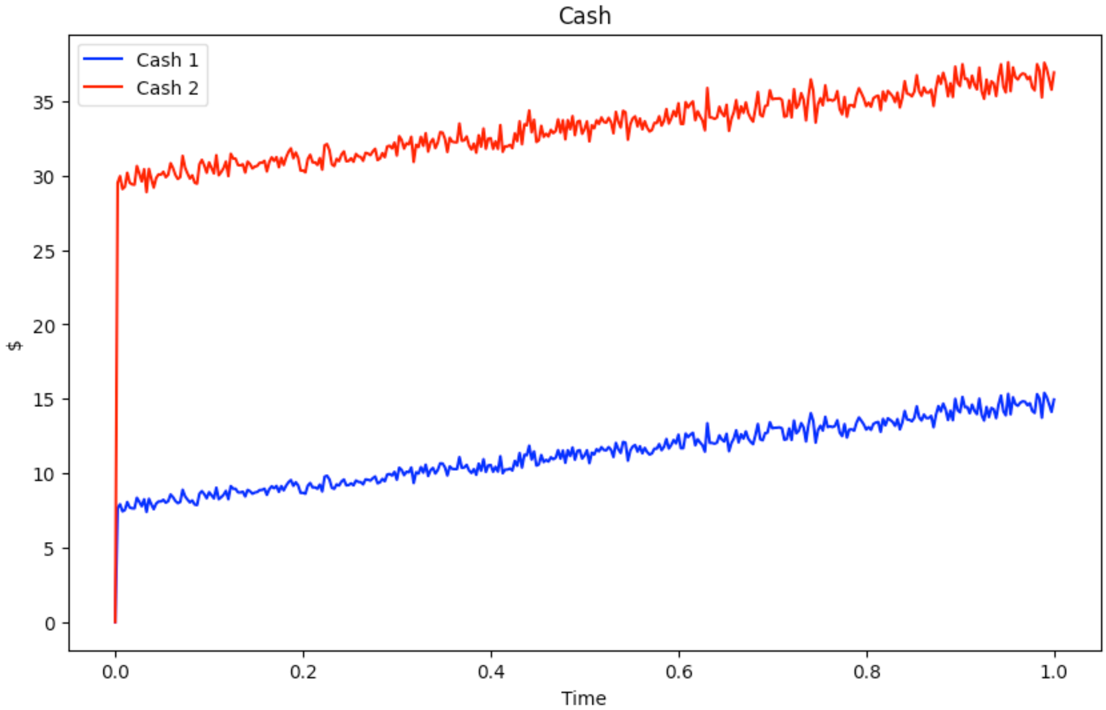

Figure 8: Dynamic cash

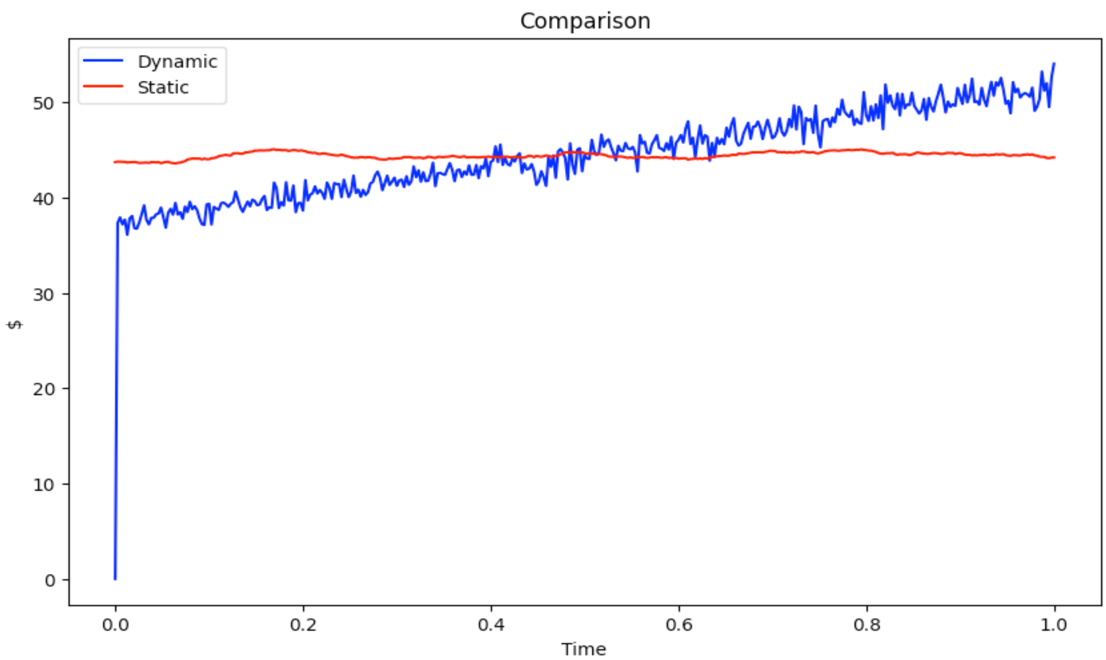

Figure 9: Return Comparison

The simulated price patterns of stock 1 and stock 2 offer important information about how well the pairs trading technique may work in a fictitious market. The strategy's main steps can be repeated using simulated data to verify its efficacy, evaluate its robustness, and comprehend the differences between dynamic and static control elements. This thorough method, which is supported by visuals, is essential for honing and verifying the trading strategy prior to implementing it in real-world trading situations.

References

Bernstein, D. and Nielsen, B. (2019). Asymptotic Theory for Cointegration Analysis When the Cointegration Rank Is Deficient. *Econometrics*, 7(1), p.6. doi:https://doi.org/10.3390/econometrics7010006.

Göncü, A. and Akyıldırım, E. (2016). Statistical Arbitrage with Pairs Trading. *International Review of Finance*, 16(2), pp.307–319. doi:https://doi.org/10.1111/irfi.12074.

Krauss, C. (2016). STATISTICAL ARBITRAGE PAIRS TRADING STRATEGIES: REVIEW AND OUTLOOK. *Journal of Economic Surveys*, 31(2), pp.513–545. doi:https://doi.org/10.1111/joes.12153.

Lintilhac, P.S. and Tourin, A. (2016). Model-based pairs trading in the bitcoin markets. *Quantitative Finance*, 17(5), pp.703–716. doi:https://doi.org/10.1080/14697688.2016.1231928.

Mori, M. and Ziobrowski, A.J. (2011). Performance of Pairs Trading Strategy in the U.S. REIT Market. *Real Estate Economics*, 39(3), pp.409–428. doi:https://doi.org/10.1111/j.1540-6229.2010.00302.x.

Patell, J.M. and Wolfson, M.A. (1984). The intraday speed of adjustment of stock prices to earnings and dividend announcements. *Journal of Financial Economics*, 13(2), pp.223–252. doi:https://doi.org/10.1016/0304-405x(84)90024-2.

Yip Chee-Yin and Hock Eam Lim (2014). The opportune time to invest in residential properties - Engle-Granger cointegration test and Granger causality test approach. *Nucleation and Atmospheric Aerosols*. doi:https://doi.org/10.1063/1.4903696.

Appendices

Phyton Code uploaded to GitHub:

https://github.com/conquerorpulkit/HFT/blob/main/AS2.ipynb