Imperial College
London

MEng Individual Project

Imperial College London

Department of Computing

Cryptocurrency Statistical Arbitrage on Decentralised Exchanges

Author:
Devam Savjani

Supervisor:
Dr. Thomas E. Lancaster

Second Marker:
Mr. Ivan Procaccini

January 26, 2023

Contents

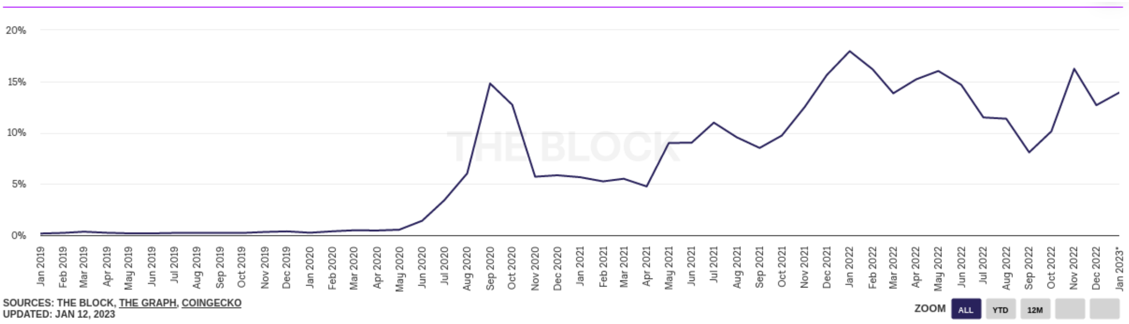
Chapter 1

Introduction

1.1 Motivation

Since the introduction of Bitcoin, a peer-to-peer payment network and cryptocurrency, there has been countless new cryptocurrencies that is used and traded. Along with this the high volatility has piqued a lot of retail investors interest with some investors gaining a high return of interest and others losing a lot of money [1]. In addition to this, larger investment institutions have also sought to gain profits from this new type of tradable asset [2]. This has lead there to be more sophisticated forms of cryptocurrency trading.

However, trading cryptocurrencies was initially very difficult for people without technical know-how, and the first recorded transaction was on 12^{th} October 2009 via a paypal transaction [3]. Since then, there has been countless other cryptocurrency exchanges have emerged. Many of which are centralized, which provide a simple interface and support for investors with little technical know-how, and others are decentralised, which work by directly interacting with the blockchain. Centralized exchanges are predominantly used as we can see in Figure 1.1, however we can also see that the volume traded in DEXes have been increasing since 2020.



SOURCES: THE BLOCK, THE GRAPH, COINGECKO
UPDATED: JAN 12, 2023

Figure 1.1: DEX to CEX Spot Trade Volume [4]

This poses the question that about which trading strategies can exploit arbitrage opportunities on decentralised exchanges. There has been some research on this topic, mainly focussing on triangular and cyclic arbitrage on DEXes such as Uniswap and SushiSwap, however there has been no research into analysing the performance of statistical arbitrage methods on decentralised exchanges.

1.2 Objectives

The objectives for this project is to design, implement and analyse the performance of numerous statistical arbitrage strategies.

1.2.1 Example of Statistical Arbitrage

Before outining the objectives, it is important to understand the process of statistical arbitrage. The process in which statistical arbitrage is executes is in 3 steps:

1. Identify a pair of assets that are cointegrated/correlated, i.e. follow the same trend and have a similar mean

2. When the prices diverge, long the undervalued asset and short the overvalued asset.

3. When the prices then revert back to it's mean, close the position.

An example of the price dynamic of one such pair can be seen in Figure 1.2.

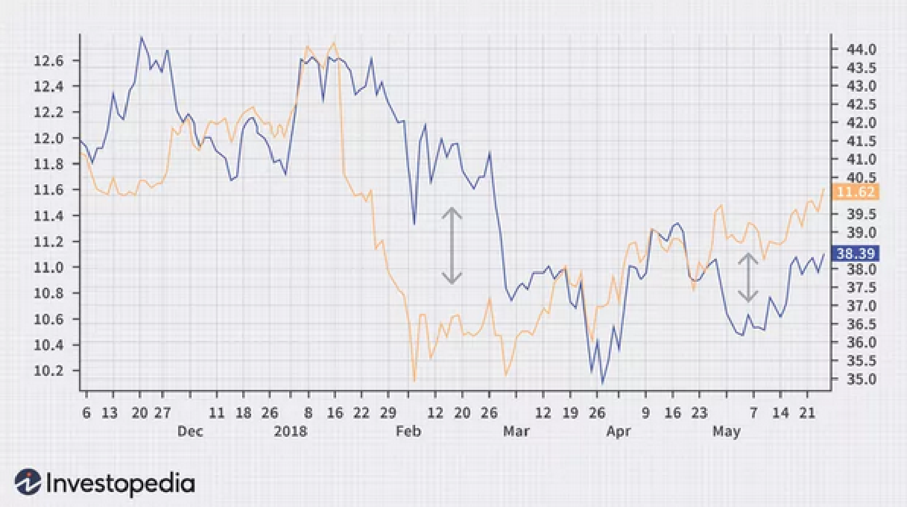


Figure 1.2: Stock Prices of General Motors and Ford Motor Company [5]

1.2.2 Objectives and Contributions

As previously mentioned, I will be implement and analyse numerous trading strategies:

- **Simple Mean Reversion Strategy**

- **Mean Reversion Strategy using the Kalman Filter to calculate the hedge ratio**

- **Machine Learning Forecasting with the Kalman Filter**

Chapter 2

Background

2.1 Cryptocurrencies

Before going delving into the financial side of the project, it is important to understand the underlying assets and the technology that drives them.

2.1.1 Blockchain

The building blocks of cryptocurrencies come from blockchain technology. Blockchain is a distributed ledger that stores data, in blocks, in a chain, comprising the data itself as well as a full transaction history [6]. Below shows a diagram of blocks in a blockchain.

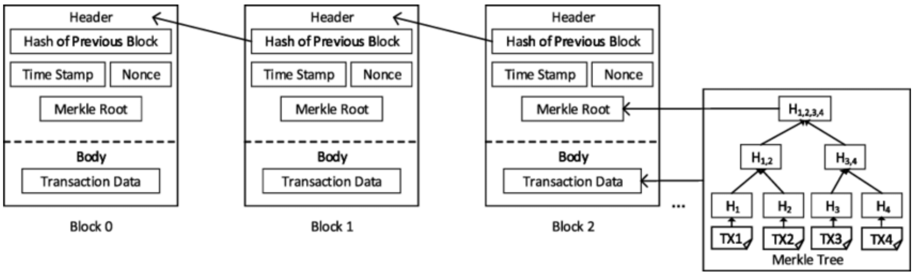


Figure 2.1: Blockchain Diagram [7]

Header, Hash of Previous Block and Timestamp

The timestamp and hashes of the block and its' preceding block are all used to ensure the ordering of blocks within a chain. Hashing the data to a fixed size, and storing it in its succeeding block makes the tampering of chains difficult as it would mean the chain deviates from its old state. In addition to this, by hashing and using Nonce, blockchain employs the Proof-of-Work algorithm to ensure correctness. The Proof-of-Work algorithm is used to confirm and add a new transaction to the chain.

Nonce

A nonce, 'Number Only used Once', is a number that is added to a hashed block to make the transaction more secure. It is randomly generated which miners use to validate a transaction. A miner first guesses a nonce and appends the guess to the hash of the current header. The miner then rehashes the value and compares this to the target hash. If the guess was correct, the miner is granted the block [8].

Merkle root

A Merkle root is also stored in each block to validate transactions efficiently, in terms of storage and searching. A Merkle tree is a tree of hashes where each leaf node is its data hash and its parent node, is the hash of their children's hashes. In storing the Merkle root, we do not need to directly store each transaction in each block, and also allows a quick search for any malicious alterations in differing blocks [9].

2.1.2 Ethereum

One of the first applications of blockchain was by Satoshi Nakamoto to create the first 'purely peer-to-peer version of electronic cash' [10]. Nakamoto's solution details the process in which a decentralised, peer-to-peer approach to verify and track transactions without a centralized institution. Since then, many other technologies derived from using blockchain as its underlying technology. One of them was proposed by Vitalik Buterin, the co-founder of Ethereum, in a whitepaper that proposed the idea of using smart contracts to create financial products and services that could operate independently of traditional financial institutions, hence decentralised finance was birthed [11].

Ethereum

Ethereum's architecture is similar to bitcoin's but has a few differences, one of which is the blockchain contains a copy of the transaction list and the most recent state. The process of how transactions are validated is below:

1. Validate the parent block
2. Validate that the current timestamp is greater than the previous timestamp
3. Check that the Ethereum concepts are valid
4. Perform Proof of Stake on the block
5. Check for errors and gas
6. Validate the final state

Proof-of-Stake (PoS) is a consensus protocol that is used by Ethereum and Bitcoin for the entire network to agree on the state of the blockchain. This provides security for malicious users to attempt to alter, add/remove transactions or maintain a second chain, the blockchain as it requires 51% of the network to agree on the alteration.

To measure how much computational effort is required to execute operations on the Ethereum network, gas is used [12]. Every block has a base fee, derived from the demand for the block space, which is burnt. Therefore, users of the network are expected to set a tip (priority fee) to reimburse miners for adding their transaction in blocks, thus the higher the tip, the greater the incentive for miners to validate the transaction. Using gas means that the Ethereum network is tolerant to spam and also has a maximum gas fee to make Ethereum tolerant to malicious code that would be used to waste resources.

Another difference between Ethereum and other cryptocurrencies is that rather than managing a distributed ledger, it uses a distributed state machine. The Ethereum Virtual Machine (EVM) defines the rules of changing states from block to block. Each node on the Ethereum blockchain contains an immutable instance of the EVM [13].

Smart Contracts

Smart contracts are programs that are self-executing contracts between buyers and sellers that deploy on the Ethereum network. It allows for the automation of a contract's execution

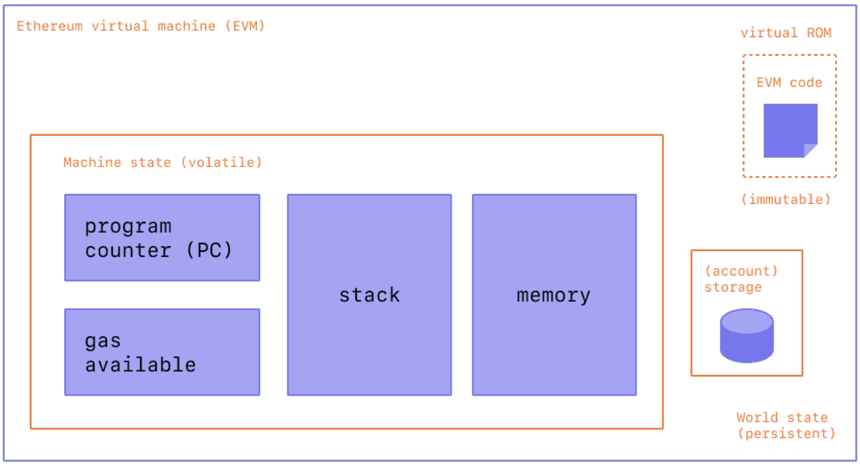


Figure 2.2: EVM components [13]

and can be used to facilitate, verify, and enforce the negotiation or performance of a contract [14, 15].

2.1.3 Decentralised Finance

One of the main applications of Ethereum and smart contracts is Decentralised Exchanges (DEXes). Before delving into DEXes it is important to understand centralized exchanges.

Centralized Exchanges

Centralized exchanges allow agents to discover and trade assets. CEXes facilitate trading between buyers and sellers by proving an online platform that manages and maintains an order book. An order book aggregates buy and sell orders and execute matching buy and sell orders. The order book and transactions are typically managed on a database as opposed to interacting with the blockchain. When trading, exchanges charge trading fees for the maker and the taker to operate the exchange and do not charge any gas fees as there is no interaction with the blockchain.

Decentralized Exchanges

In contrast, DEXes utilize blockchain technology and smart contracts to execute trades thus providing a high level of determinism. These trades are executed on the blockchain via smart contracts and on-chain transactions. There are two types of DEX, order book DEXes and Automated Market Makers (AMMs). An order book DEX is less common and is similar to CEXes however, the order book is stored on the blockchain rather than on a central database. This means each order placed requires the order book to be posted on the blockchain at each transaction. Automated Market Makers are more common and provide instant liquidity by using liquidity pools so that users can swap their tokens for a price that is determined by the portions within the liquidity pool [16]. DEXes have multiple pros including lower transaction fees, privacy, diversity and trustless transactions but they also have their drawbacks such as scalability and poor liquidity a lot of the DEXes are quite new [17].

2.2 Arbitrage

Arbitrage is the process in which a trader simultaneously buys and sells an asset to take advantage of a market inefficiency [18]. Arbitrage is also possible in other types of securities by finding price inefficiencies in the prices of options, forward contracts and other exotics.

Sources have shown that the word "*Arbitrage*" has been used as early as the Renaissance era when surviving documents showed a large number of bills being exchanged [19]. There has also been some evidence to suggest that arbitrage was used as early as the Greek and Roman eras. Objects such as Sumerian cuneiform tablets show the trade of ancient bills however we cannot come to strong conclusions about this. Early forms of arbitrage would likely have been purchasing a commodity then transporting them to a foreign land and selling them at a higher price. This type of arbitrage is called commodity arbitrage and is still applicable today. With the example above, transporting the goods takes a significant amount of to the merchant, or trader, which could cause variations in the price, however, in the modern day this has been reduced and with electronic exchanges, this time to buy and sell is very small. This means inefficiencies in the market, where a trader can profit purely by buying and selling, should not exist. This is called the "Law of One Price". The "Law of One Price" states that every identical commodity or asset should have the same price regardless of exchange or location, given there are no transaction costs, no transportation costs, no legal restrictions, the exchange rates are the same and no market manipulation occurs [20]. This is because if this were not the case, an arbitrage opportunity would arise and someone would take advantage of the scenario causing the prices on both markets to converge due to the market forces. In the real world arbitrage opportunities are tremendously common, thus allowing a risk-free investment [21, 22].

There are countless types of arbitrage such as spatial arbitrage, which profits off of different prices on exchanges in different locations, temporal arbitrage, which takes advantage of price differences at different times, risk arbitrage, which profits from perceived discrepancies in their risk-return profiles and finally market arbitrage which takes advantages of different prices on different exchanges/markets. Statistical methods include pairs trading, which involves buying and selling assets that are believed to be mispriced relative to one another, momentum trading, which identifies if assets have a strong momentum (either up or down) and profiting off of that, and finally, algorithmic trading which uses algorithms to analyze data and trades based on statistical analysis. This project shows how these opportunities can be exploited both in a pure manner as well as using statistical methods.

2.3 Pure Arbitrage Techniques

To better understand the project and to be able to research something new and novel it is important to understand the current state of art, i.e. previous research on the topic. Research into cryptocurrency arbitrage is still in its infancy and previous research has mainly focussed on the economics of cryptocurrencies, i.e. miner/trader behaviour and influence of cryptocurrency trading [23, 24, 25, 26, 27, 28, 29]. Furthermore, there has been very limited research comparing statistical strategies and pure methods of arbitrage of cryptocurrencies. Despite this, there has been plentiful research on arbitrage as a whole as it is immensely profitable, as a result of this people/institutions tends to keep their newly found research secret. Of the published research, I have looked into the arbitrage techniques that are used. As arbitrage can is highly profitable, it can be found in countless types of assets, such as options, stocks, bonds and many other types of products. Research into all types of products exist going into the theory and practical aspects of each [30, 31]. The most similar type of asset class to cryptocurrencies is fiat currencies, such as the US Dollar and the Great British Pound Sterling. The research in arbitrage in foreign exchanges shows that using a triangular/cyclic arbitrage is highly profitable and effective [32, 33, 34].

Although this project aims to focus on statistical arbitrage techniques, it is still important to look at research that looks into purer forms of arbitrage, i.e. triangular and cyclic arbitrage.

As previously mentioned, research into this topic is still in its infancy thus which means a very thin slice of exploration on the subject matter. The majority of the research has been into the arbitrage on centralized exchanges, [35, 36, 37]. Cristian Pauna investigates and implements an arbitrage strategy in [37]. The paper details the technical details of arbitrage trading from the data and the system architecture used. Pauna finds complications such as requesting data from multiple exchanges, converting the data such that it is homogeneous and also managing server load. Pauna presents the architecture such that the servers request data from the necessary exchanges, aggregating prices in a relational database which then triggers a server that is used to generate trading signals.

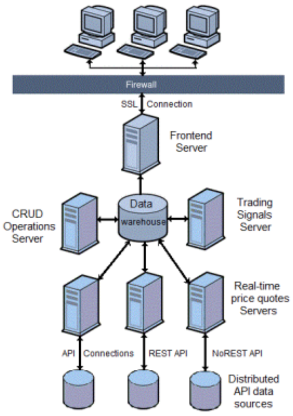


Figure 2.3: Arbitrage system architecture [37]

As previously mentioned triangular and cyclic arbitrage is one of the most used and purest forms of arbitrage to implement and analyse, [38] explores triangular arbitrage on decentralised exchanges. Algorithm 1 is the algorithm used to find the most profitable arbitrage route on a particular platform, once this is calculated, it is compared with other routes on other platforms. Initially, the system converts the base token into another token and converts it back into the base token, using only one token is used as a middle route, then using the algorithm below, increases the number of middle tokens.

Algorithm 1 Maximum Profit Route Searching (R)

```
Input: $T$ (token list), $P$ (price graph), $n$ (current route)
  for $i = 1, ..., T$ do
      $r = get\_profit(n + i)$
      for $j = 1, ..., P[i]$ do
          $p = max(r, R(T, P, n_j))$
      end for
  end for
  return $p$
```

On evaluating the performance of the strategy on differing platforms depended on three main features of each exchange:

1. Portion size - Depending on how much the “trader” invested revenues differed and with the larger portion size, the revenue decreases as the token pair prices are adjusted based on supply/demand.

2. Transaction fees - Each exchange has its own transaction fee.

3. Other considerations such as price slippage - Exchanges have different liquidity levels which depend on the usage and liquidity providers that the exchange employs.

It is found that using this strategy out of the exchanges; Uniswap, 1inch, Kyberswap and Bancor. 1 inch was the only exchange that generated a profit whereas the others lose money.

Another paper that implemented and evaluated a cyclic arbitrage opportunity is [39]. The research consists of proposing a theoretical arbitrage model and further evaluation of real transactional data. The arbitrage model used is simple to understand, as it searches for a cyclic transaction between n tokens, $A_1, A_2, ..., A_n$ is a sequence of n trades:

Trade 1: Exchange δ_1 of A_1 to δ_2 of A_2

Trade 2: Exchange δ_2 of A_2 to δ_3 of A_3

...

Trade n: Exchange δ_n of A_n to δ'_1 of A_1

It is important to note that $\delta_i = \delta_{i+1}$, i.e. the output of trade is equivalent to the input of the next. The revenues within a cycle are defined as $\delta_{i+1} - \delta_i$, and the overall profit is $\delta'_1 - \delta_1$. This is not as simple as the revenues depend on how liquid the exchange is, thus the liquidity pools of each possible trading pair are hugely important. Therefore, the paper proposes a theorem, below:

Theorem 1 *For a given cycle $A_1 \rightarrow A_2 \rightarrow \cdots \rightarrow A_n \rightarrow A_1$ with n tokens, there exists an arbitrage opportunity for the cyclic transaction if the product of exchange rates $\frac{a_{2,1}a_{3,2}\cdots a_{1,n}}{a_{1,2}a_{2,3}\cdots a_{n,1}} > \frac{1}{r_1^n r_2^n}$ where $a_{i,j}$ denotes the liquidity of token A_i in the liquidity pool with token A_j. [39]*

In addition to the theorem, to obtain an optimal strategy we need to compute the optimal trading volume of a cycle, $A_1 \rightarrow A_2 \rightarrow \cdots \rightarrow A_n \rightarrow A_1$. The paper proposes the optimal trading volume to be $\delta_a^{op} = \frac{\sqrt{r_1 r_2 a' a} - a}{r_1}$ where $a = \frac{a'_{1,n} a_{n,1}}{a_{n,1} + r_1 r_2 a'_{n,1}}$ and $a' = \frac{r_1 r_2 a'_{1,n} a_{n,1}}{a_{n,1} + r_1 r_2 a'_{n,1}}$. Thus to calculate such arbitrage opportunities knowing the liquidity of tokens in other tokens' liquidity pools, algorithm 2 infers the direction and volumes to trade to get the optimal revenue.

Algorithm 2 Computing the equivalent liquidity of the cycle

```
$a'_{1,n} \leftarrow a_{1,2}$
$a'_{n,1} \leftarrow a_{2,1}$
for $i$ from 2 to $n-1$ do
    $a'_{1,n} \leftarrow \frac{a'_{1,n} a_{i,i+1}}{a_{i,i+1} + r_1 r_2 a'_{n,1}}$
    $a'_{n,1} \leftarrow \frac{r_1 r_2 a'_{1,n} a_{i+1,i}}{a_{i,i+1} + r_1 r_2 a'_{n,1}}$
end for
```

After analysis, it is found that between 4th May 2020 to 15th April 2021, there were countless exploitable arbitrage opportunities which grew to 1,750 in the 11 months that it was tested. Only cycles with length 3 were experimented with and only cycles including ETH as 80% of the liquidity pools on Uniswap include ETH and another cryptocurrency [40]. Furthermore, it is found that 287,241 of the 292,606 arbitrages executed started with ETH, and 85% of the arbitrages used a cycle of length 3. The total revenue of the cyclic arbitrage was 34,429 ETH. However, gas fees account for 24.6% of the total revenue leaving an approximate 25,971 ETH profit.

The paper then delves into the implementation of the smart contract, the paper explored how both *sequential* and *atomic* implementations would affect the revenue and execution of the contracts. It was found that 52.3% of the arbitrages that were executed sequentially generated a loss, likely due to the fact when one submits n orders, the n blockchain transactions are executed sequentially, meaning some external transactions can be inserted between these transactions. Thus using atomic transactions avoids this issue of external transactions does not affect the market price that may affect the outcome of the arbitrage.

Furthermore, the authors of the paper also investigated the performance differences between using private smart contracts and public contracts. Deploying a smart contract that calls Uniswap functions, i.e. a private smart contract, is intuitively better and achieves a higher success rate of a lower bound of 52% and a higher bound of 90% in comparison to calling a public Uniswap smart contract which has a success rate of 27.3%. Overall the paper provides an insightful look into cyclic arbitrage in DEXes and highlights important decisions made such as liquidity calculations and smart contracts while comparing the performance of different options available.

2.4 Statistical Arbitrage Techniques

As mentioned previously mentioned and the subject of the project is to optimize statistical arbitrage methods to be able to compete with a purer form of arbitrage, i.e. cyclic arbitrage. As previously mentioned there are many methods of stat arb, pairs trading, momentum trading and algorithmic trading. Within these methods there are countless strategies to adopt and profit from, thus to limit the scope, this project I will be investigating strategies within pair trading. Research within Pair trading has been vast with many streams of approaches emerging; distance approach, cointegration approach, time-series approach, stochastic approach and some others, including using machine learning [41]. However, for this project, we will only look at cointegration/co-correlation approaches.

2.4.1 Mean Reversion

The cointegration approach follows three key steps. The first is the selection of pairs based on similarity measures, the next is assessing the tradability and finally, thresholds are set for trading. The spread is defined as

$$\varepsilon_{ij,t} = P_{i,t} + \gamma P_{j,t}$$

where $P_{i,t}$ and $P_{j,t}$ denote the $I(1)$ non-stationary price processes of the assets i and j, γ is the cointegration coefficient, also referred to in literature as the hedge ratio. $\varepsilon_{ij,t}$ is the linear combination of the non-stationary prices and is $I(0)$ stationary and hence mean-reverting, note that stationary processes are those of which have a constant mean. Rad's implementation of this approach on stocks results in a 0.83% return before considering transaction costs [42]. Another paper, [43], looked into setting the thresholds and setting a minimum profit, MP_{ij,t_c}:

$$MP_{ij,t_c} = \frac{n(\varepsilon_{ij,t_0} - \varepsilon_{ij,t_c})}{\mid \gamma \mid}$$

Where t_0 and t_c are the opening and closing times, n is the volume longed of asset j.

2.4.2 Optimal Portfolio Design for Mean Reversion

There has been further research into optimizing mean reversion, one of which was to use the successive convex approximation method on the mean reverting portfolio design [44]. The paper initially proposes the mean reversion portfolio:

- For each asset, the price at time t is denoted as p_t and its correspondeing log-price $y_t \triangleq log(p_t)$, its vector form of M assets $\mathbf{y}_t \triangleq \left[y_{1,t}, \ldots, y_{M,t}\right]^T$.
- The log-price spread is given by $y_t \triangleq \beta^T \mathbf{y}_t$, where $\beta \triangleq \left[\beta_1, \ldots, \beta_M\right]^T$ denotes the hedge ratios.
- The cointegration space with N relations is defined by $\mathbf{B} \triangleq \left[\beta_1, \ldots, \beta_N\right]$, thus the N spreads are $s_t \triangleq \mathbf{B}^T \mathbf{y}_t$.
- For these N spreads, the portfolio weight matrix is denoted as $\mathbf{w} \triangleq \left[w_1, \ldots, w_N\right]^T$.
- The auto-covariance matrix for the spreads s_t is defined as $M_i \triangleq Cov(s_t, s_{t+i}) = \mathbb{E}\left[\left(s_t - \mathbb{E}\left[s_t\right]\right)\left(s_{t+i} - \mathbb{E}\left[s_{t+i}\right]\right)^T\right]$

Now that we have defined everything required, we can now formalize the problem. The general problem of mean reversion portfolio design problem is formalized by:

$$\begin{aligned} &\underset{\mathbf{w}}{\text{minimize}} && F(\mathbf{w}) \triangleq U(\mathbf{w}) + \mu V(\mathbf{w}) + \gamma S(\mathbf{w}) \\ &\text{subject to} && \mathbf{w} \in \left\{\mathbf{w} \mid \left\|\mathbf{B}\mathbf{w}\right\|_0 \leq L\right\}, \qquad \text{where } L \text{ is the total leveraged investment} \end{aligned}$$

- μ defines the trade-off between the mean reversion measure and the variance preference.

- γ defines the regularization parameter of how sparse we would like the cointegration space to be.

Where the Mean Reversion term:

$$U(\mathbf{w}) \triangleq \xi \frac{\mathbf{w}^T \mathbf{H} \mathbf{w}}{\mathbf{w}^T \mathbf{M}_0 \mathbf{w}} + \zeta \left(\frac{\mathbf{w}^T \mathbf{M}_1 \mathbf{w}}{\mathbf{w}^T \mathbf{M}_0 \mathbf{w}}\right)^2 + \eta \sum_{i=2}^{p} \left(\frac{\mathbf{w}^T \mathbf{M}_i \mathbf{w}}{\mathbf{w}^T \mathbf{M}_0 \mathbf{w}}\right)^2$$

And the variance term:

$$V(\mathbf{w}) \triangleq \begin{cases} 1/\mathbf{w}^T \mathbf{M}_0 \mathbf{w} & \text{VarInv}(\mathbf{w}) \\ 1/\sqrt{\mathbf{w}^T \mathbf{M}_0 \mathbf{w}} & \text{StdInv}(\mathbf{w}) \\ -\mathbf{w}^T \mathbf{M}_0 \mathbf{w} & \text{VarNeg}(\mathbf{w}) \\ -\sqrt{\mathbf{w}^T \mathbf{M}_0 \mathbf{w}} & \text{StdNeg}(\mathbf{w}) \end{cases}$$

The variance term can be represented in any of the four forms.

And the asset selection term:

$$S(\mathbf{w}) \triangleq \left\|\mathbf{B}\mathbf{w}\right\|_0 = \sum_{m=1}^{M} sgn(\mid \left[\mathbf{B}\mathbf{w}\right]_m \mid)$$

This asset selection criterion is not necessary however as trading incurs a cost, selecting all of the assets is costly, thus selecting a subset of assets to trade is more profitable. To formalize this goal, we would like to minimize the cointegration space thus we use the ℓ_0 norm.

The paper then goes on to solve the optimization problem using the successive convex approximation (SCA) method [45]. The SCA method takes an optimization problem in the form of:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} && f(\mathbf{x}) \\ &\text{subject to} && \mathbf{x} \in \mathcal{X} \end{aligned}$$

Where $\mathcal{X} \subseteq \mathbb{R}^N$ is convex and $f(\mathbf{x})$ is non-convex. The SCA method involves starting at an initial point $\mathbf{x}^{(0)}$ and solving a series of subproblems of surrogate funtions $\tilde{f}(\mathbf{x};\mathbf{x}^{(k)})$ over the set $\mathcal{X}$. The sequence $\{\mathbf{x}^{(k)}\}$ is generated by:

$$\begin{cases} \hat{\mathbf{x}}^{(k+1)} = \underset{\mathbf{x}\in\mathcal{X}}{argmin} \tilde{f}(\mathbf{x};\mathbf{x}^{(k)}) \\ \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \gamma^{(k)}(\hat{\mathbf{x}}^{(k+1)} - \mathbf{x}^{(k)}) \end{cases}$$

The first step is to generate a descent direction and then update the variable with a step size of $\gamma^{(k)}$. After applying this method to the MRP problem and further analysis of the paper, the following algorithm is proposed and used to solve the MRP design problem:

Algorithm 3 SCA-Based Algorithm for The Optimal MRP Design Problem

```
Require: $\mathbf{H}, \mathbf{M}_i, \mu, \gamma, \mathbf{B}, L$ and $\tau$
1: Set $k = 0, \gamma^{(0)}$ and $\mathbf{w}^{(0)}$
2: repeat
3:     Compute $\mathbf{A}^{(k)}$ and $\mathbf{b}^{(k)}$
4:     $\hat{\mathbf{w}}^{(k+1)} = \underset{\mathbf{w}\in\mathcal{W}}{argmin} \mathbf{w}^T\mathbf{A}^{(k)}\mathbf{w} + \mathbf{b}^{(k)T}\mathbf{w}$
5:     $\mathbf{w}^{(k+1)} = \mathbf{w}^{(k)} + \gamma^{(k)}(\hat{\mathbf{w}}^{(k+1)} - \mathbf{w}^{(k)})$
6:     $k \leftarrow k + 1$
7: until convergence
```

However, 4 line is a convex problem and has no closed-form solution thus to solve this subproblem using the ADMM method, this is done by introducing an auxiliary variable $\mathbf{z} = \mathbf{Bw}$.

$$\begin{aligned} \underset{\mathbf{x},\mathbf{z}}{\text{minimize}} \quad & \mathbf{w}^T\mathbf{A}\mathbf{w} + \mathbf{b}^T\mathbf{w} \\ \text{subject to} \quad & \left\|\mathbf{z}\right\|_1 \leq B, \mathbf{Bw} - \mathbf{z} = \mathbf{0} \end{aligned}$$

This is then summarized into Algorithm 4:

Algorithm 4 An ADMM-Based Algorithm for Problem on line 4 in Algorithm 3

```
Require: $\mathbf{A}, \mathbf{b}, \mathbf{B}, B, \rho$
Set $\mathbf{w}^{(0)}, \mathbf{z}^{(0)}, \mathbf{u}^{(0)}$ and $k = 0$
repeat
    $\mathbf{w}^{(k+1)} = -(2\mathbf{A} + \rho\mathbf{B}^T\mathbf{B})^{-1}(\mathbf{b} + \rho\mathbf{B}^T(\mathbf{u}^{(k)} - \mathbf{z}^{(k)}))$
    $\mathbf{h}^{(k)} = \mathbf{Bw}^{(k+1)} + \mathbf{u}^{(k)}$
    $\mathbf{z}^{(k+1)} = \Pi_{\mathcal{C}}(\mathbf{h}^{(k)})$
    $\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \mathbf{Bw}^{(k+1)} - \mathbf{z}^{(k+1)}$
    $k \leftarrow k + 1$
until convergence
```

After all of this analysis, the authors of the paper, [46, 44], ran simulations on real data comparing underlying spread. It found that it resulted in consistent profits as shown in figure 2.4:

Overall, this research successfully formulizes, solves the optimization problem mathematically, and goes further to implement the algorithms to solve the problem programmatically. In addition, the author compares the implementation with other benchmark algorithms, showing that it results in a greater P&L and Sharpe ratio.

2.4.3 Statistical Arbitrage using the Kalman Filter

Another method that is used in statistical arbitrage is using the Kalman Filter. Recall the equation for spread:

$$\varepsilon_{ij,t} = P_{i,t} + \gamma P_{j,t}$$

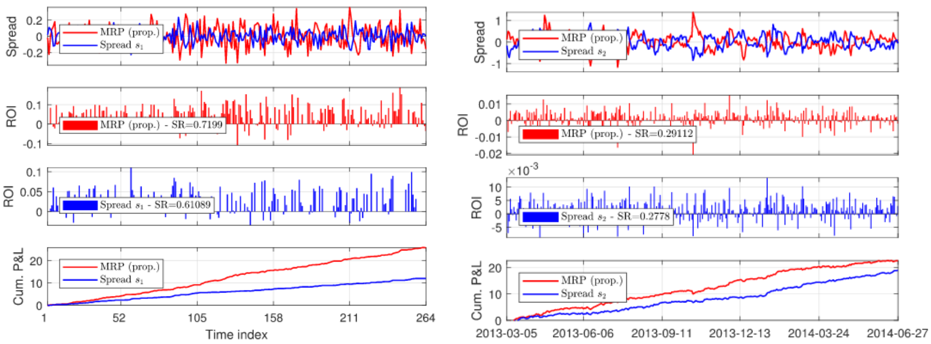


Figure 2.4: A mean-reversion trading based on real data [44]

The Kalman filter is a recursive algorithm for estimating the state of data that is usually very noisy and thus needs to be filtered [47]. This makes it very useful to estimate the hedge ratio γ. Initially, a book by Vidyamurthy discusses best practices for choosing cointegrated equities and found that the Kalman filter was found optimal when the state-space and observation equations are linear and the noise is Gaussian [48]. Since then there have been many extensions of the filter such as the Extended Kalman Filter (EKF) and Unscented KF aimed to handle when the state-space and observation equations are non-linear and the noise is not Gaussian.

The Kalman Filter works in 2 phases, prediction and update. The prediction phase is as follows

$$\hat{\mathbf{x}}_k = \mathbf{F}_k \hat{\mathbf{x}}_{k-1} + \mathbf{B}_k \vec{\mathbf{u}}_k + \mathbf{w}_k$$

$$\mathbf{P}_k = \mathbf{F}_k \mathbf{P}_{k-1} \mathbf{F}_k^T + \mathbf{Q}_k$$

Where $\hat{\mathbf{x}}_k$ is the new best estimate (prediction) that is derived from $\hat{\mathbf{x}}_{k-1}$, the previous estimate and the prediction function $\mathbf{F}_k$. $\vec{\mathbf{u}}_k$ is the correction term, called the control vector, that is used when it is known that there are external influences in combination with $\mathbf{B}_k$ which is called the control matrix. In addition to this, the new uncertainty (covariance matrix), $\mathbf{P}_k$, is calculated using the previous uncertainty and additional uncertainty from the environment, $\mathbf{Q}_k$, $\mathbf{w}_k$ is called the state noise. The update is as follows

$$\hat{\mathbf{x}}'_k = \hat{\mathbf{x}}_k + \mathbf{K}'(\vec{\mathbf{z}}_k - \mathbf{H}_k \hat{\mathbf{x}}_k)$$

$$\mathbf{P}'_k = \mathbf{P}_k - \mathbf{K}' \mathbf{H}_k \mathbf{P}_k$$

$$\mathbf{K}' = \mathbf{P}_k \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1}$$

Where $\mathbf{K}'$ is defined as the Kalman gain, $\mathbf{H}_k$ is the measurement matrix, $\vec{\mathbf{z}}_k$ is mean of the observed values, which is also calculated by $\vec{\mathbf{z}}_k = \mathbf{H}_k \hat{\mathbf{x}}_k + \mathbf{v}_k$ where $\mathbf{v}_k$ is the measurement noise, and $\mathbf{R}_k$ is the covariance of the uncertainty of the observed values [49].

A paper that investigated the use of the Kalman Filter on ETFs found that the strategy it employed worked well for in-sample data points and worse, but still profitable, results for out-of-sample data. The paper adapted the Kalman Filter to be able to use it for pairs trading to the following:

$$\mathbf{y}_t = \mathbf{x}_t \beta_t + \epsilon_t$$

$$\beta_t = \mathbf{I} \beta_{t-1} + \omega_t$$

Then calculating the Kalman Gain:

$$\text{Kalman Gain} = \frac{\text{Error in the estimate}}{\text{Error in the estimate} + \text{Error in the measurement}}$$

Then to calculate the estimate:

$$\text{Estimate}_t = \text{Estimate}_{t-1} + \text{Kalman Gain} \times (\text{Measurement - Estimate}_{t-1})$$

And finally, calculating the new error:

$$E_{\text{estimate}_t} = \frac{E_{\text{measurement}} \times E_{\text{estimate}_{t-1}}}{E_{\text{measurement}} + E_{\text{estimate}_{t-1}}}$$

$$E_{\text{estimate}_t} = E_{\text{estimate}_{t-1}} \times (1 - \text{Kalman Gain})$$

The paper later hypothesises the causes for the disappointing results to the "pairs trading strategies have gained widespread acceptance thus making profitability much more elusive", however, the author fails to find evidence or provide sufficient evidence to justify the claim [50].

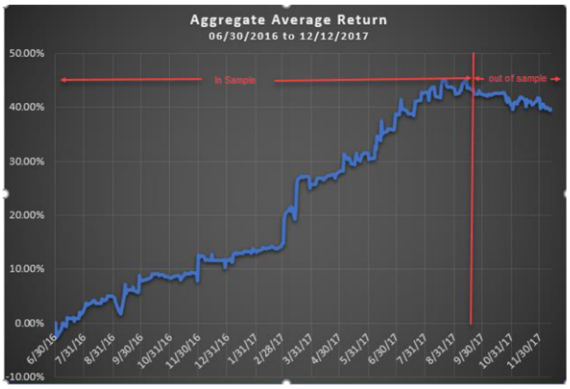


Figure 2.5: Aggregate average return of using the Kalman filter for pairs trading on ETFs [50]

Another paper used the combination of the Kalman Filter and Machine Learning, more specifically Extreme Learning Machine and Support Vector Regression (SVR) to build a statistical arbitrage strategy on the Brazilian Stock Exchange. The strategies can simply be explained as using SVR and ELM to forecast returns and using the Kalman Filter to improve the forecast.

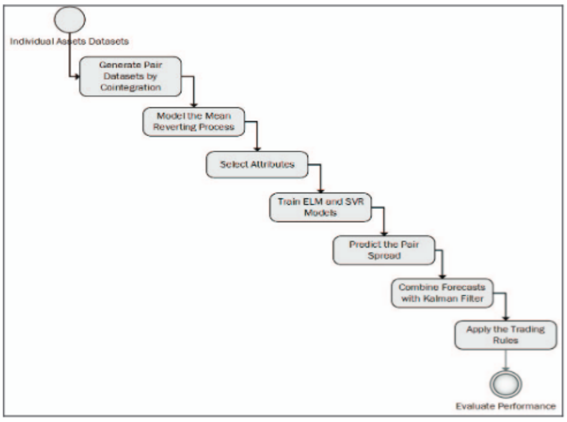

Individual Assets Datasets
Generate Pair Datasets by Cointegration
Model the Mean Reverting Process
Select Attributes
Train ELM and SVR Models
Predict the Pair Spread
Combine Forecasts with Kalman Filter
Apply the Trading Rules
Evaluate Performance

Figure 2.6: Visualisation of the trading strategy used in [51]

The paper also compares methods, such as LASSO, BMA, and GRR, to benchmark the performance of the Kalman Filter. The research found that using simply ELM and SVR forecasts results in a return of 20.19% and 21.32% respectively for out-of-sample data points and using a combination with the Kalman Filter gives a return of 26.13% for out-of-sample data points. The full results can be seen below in Figure 2.7. In addition to this, it can be seen that the volatility of the return also decreases which is ideal for investment managers.

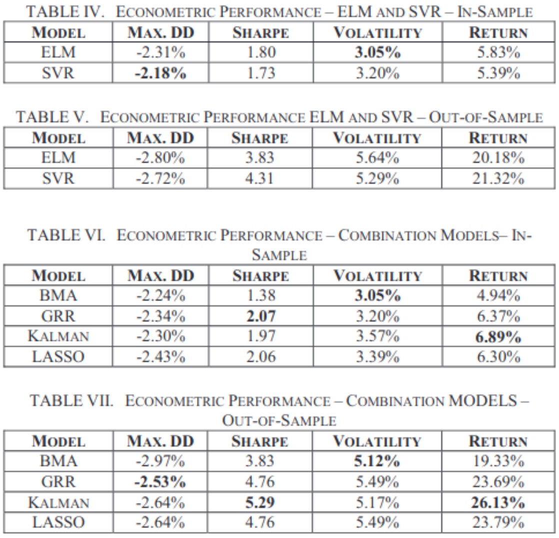


TABLE IV. ECONOMETRIC PERFORMANCE – ELM AND SVR – IN-SAMPLE

MODEL	MAX. DD	SHARPE	VOLATILITY	RETURN
ELM	-2.31%	1.80	**3.05%**	5.83%
SVR	**-2.18%**	1.73	3.20%	5.39%

TABLE V. ECONOMETRIC PERFORMANCE ELM AND SVR – OUT-OF-SAMPLE

MODEL	MAX. DD	SHARPE	VOLATILITY	RETURN
ELM	-2.80%	3.83	5.64%	20.18%
SVR	-2.72%	4.31	5.29%	21.32%

TABLE VI. ECONOMETRIC PERFORMANCE – COMBINATION MODELS– IN-SAMPLE

MODEL	MAX. DD	SHARPE	VOLATILITY	RETURN
BMA	-2.24%	1.38	**3.05%**	4.94%
GRR	-2.34%	**2.07**	3.20%	6.37%
KALMAN	-2.30%	1.97	3.57%	**6.89%**
LASSO	-2.43%	2.06	3.39%	6.30%

TABLE VII. ECONOMETRIC PERFORMANCE – COMBINATION MODELS – OUT-OF-SAMPLE

MODEL	MAX. DD	SHARPE	VOLATILITY	RETURN
BMA	-2.97%	3.83	**5.12%**	19.33%
GRR	**-2.53%**	4.76	5.49%	23.69%
KALMAN	-2.64%	**5.29**	5.17%	**26.13%**
LASSO	-2.64%	4.76	5.49%	23.79%

Figure 2.7: Econometric results [51]

Other papers/articles such as [52, 53, 54] have designed, compared and analysed other statistical arbitrage techniques using Machine Learning algorithms and revealed that some algorithms are profitable. The majority of research on machine learning trading strategies has been on assets such as stocks on centralized exchanges. The little research that has been done on statistical arbitrage on cryptocurrencies has all been on analysing arbitrage on centralized exchanges and not decentralised exchanges. One of the research projects that analysed machine learning methods of statistical arbitrage on cryptocurrencies on a centralized exchange, compared a logistic regression approach with a random forest approach [54].

2.4.4 Analysis on Cryptocurrency Arbitrage on Centralized Exchanges

Although the research in the papers previously mentioned does not investigate the cointegration approach on cryptocurrencies, the takeaways are the mathematical fundamentals that are used in statistical arbitrage. Kristoufek and Bouri researched the sources of stat. arb. of bitcoin in multiple centralized exchanges. The Grey correlation is built on top of the Grey system theory [55], and can capture non-linear correlations without assuming a Gaussian distribution, thus using the Grey correlation provides a more robust metric to understand correlations between both series. The Grey correlation $\gamma(X_0, X_i)$ is defined with two steps:

1. $\gamma(x_0(k), x_i(k)) = \frac{min_i min_k |x_0(k)-x_i(k)| + \varepsilon max_i max_k |x_0(k)-x_i(k)|}{|x_0(k)-x_i(k)| + \varepsilon max_i max_k |x_0(k)-x_i(k)|}$

2. $\gamma(X_0, X_i) = \frac{1}{n}\sum_{i=1}^{n} \gamma(x_0(k), x_i(k))$

With $\varepsilon \in [0, 1]$, the standard is set to $\varepsilon = 0.5$.

The DCC-GARCH(1,1), [56], model is also used to obtain conditional correlations for Bitcoin exchanges. The model was designed to use a combination of parameters such as the standard deviation of Bitcoin returns, traded volume, the volume of on-chain transactions, fees paid to miners, the ratio of current price and recent price history and internet hype/trends.

Upon analysis of Grey and DCC-GARCH(1,1) correlations, it is found that the DCC correlations show a little variability whereas the Grey correlations are a lot more variable ranging from 0.29 to 1. In addition, the paper then further investigates these sources and finds that these opportunities are introduced when there is a large number of inter-exchange transfer requests, i.e. the network is congested, and high price volatility. In contrast, the high volume of exchanges and on-chain activity cause the arbitrage opportunities to decrease. This paper finds and explains these sources of statistical arbitrage however does not implement or devise an algorithm that uses statistical arbitrage to generate a profit from price discrepancies of Bitcoin on different exchanges.

A paper that investigates statistical arbitrage on multiple cryptocurrencies is [57]. The authors of this paper analysed co-movements and cointegration of different cryptocurrencies on a centralized exchange using Augmented Dickey-Fuller (ADF) and Kwiatkowsky-Phillips-Schmidt-Shin (KPSS), Ljung-Box autocorrelation tests on both stationary forms ($I(0)$) and the original form ($I(1)$). The paper then develops a dynamic factor model based on the assumption that the price dynamics of cryptocurrencies are driven by Bitcoin [58], this is then evidenced by similar paths found in cryptocurrencies shown in Figure 2.8.

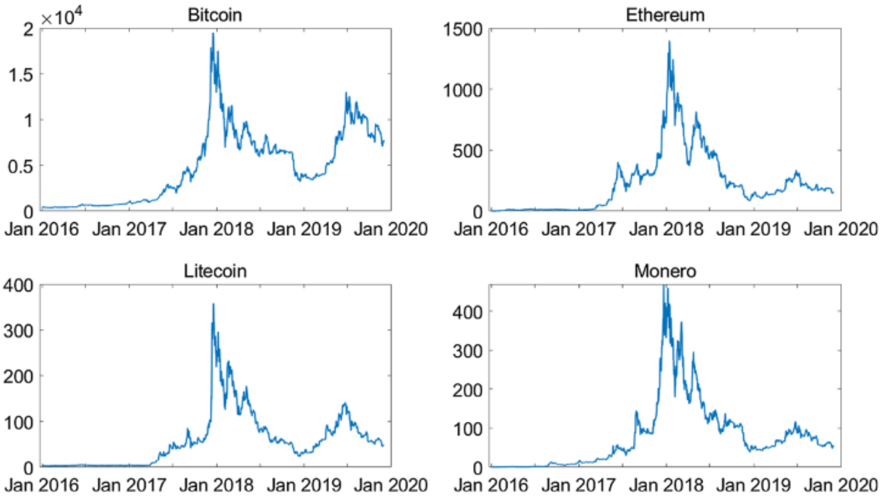


Figure 2.8: Price behaviour of Bitcoin, Ethereum, Litecoin, Monero [57]

For simplicity the authors set the number of hidden factors to 2 and upon analysis f_1 is a $I(1)$ process and the second factor f_2 is a stationary process that is independent of f_1. It is also found after overlaying f_1 with the price of Bitcoin, that the first factor strongly correlates with the price of Bitcoin.

The paper then uses this model to build an investment strategy, using forecasting using the estimated parameters:

$$\hat{p}_{i,\tau+1} = \mathbb{E}_\tau(p_{i,\tau+1}) = \hat{\alpha}_i + \hat{\beta}_{i1}\mathbb{E}_\tau(f_{1,\tau+1}) + \hat{\beta}_{i2}\mathbb{E}_\tau(f_{2,\tau+1})$$

Where

$$f_{1,t} = \lambda_1 f_{1,t-1} + \eta_{1,t}$$

$$f_{2,t} = \lambda_2 f_{2,t-1} + \eta_{2,t}$$

The expected gains one day ahead are given by:

$$g_{\tau+1} = \mathbb{E}_\tau[v_{\tau+1}] = \sum_{i=1}^{\lfloor I/2 \rfloor} \hat{p}^{(i)}_{\tau+1} - \sum_{i=\lceil I/2 \rceil+1}^{I} \hat{p}^{(i)}_{\tau+1}$$

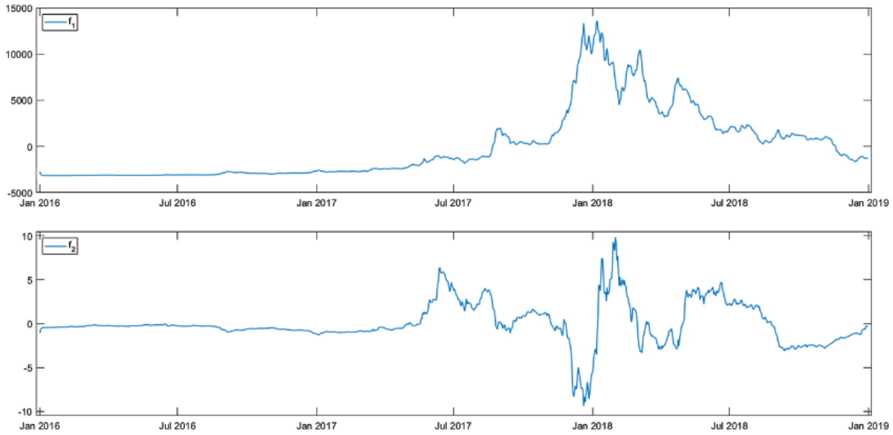


Figure 2.9: Hidden factors f_1 and f_2 from Jan 2016 to Dec 2018 [57]

Using this and a threshold which is calculated by the combination of the current price and standard deviation of the trading position value:

- if $g_{\mathcal{T}+1} > v_{\mathcal{T}} + c\sigma^v_{\mathcal{T}}$, go long
- if $g_{\mathcal{T}+1} < v_{\mathcal{T}} - c\sigma^v_{\mathcal{T}}$, go short
- if $v_{\mathcal{T}} - c\sigma^v_{\mathcal{T}} \leq g_{\mathcal{T}+1} \geq v_{\mathcal{T}} + c\sigma^v_{\mathcal{T}}$, no trade

The researchers of the paper evaluated their trading strategy for 334 days and a moving window of 3 years, 1096 observations, every day to estimate the parameters for the dynamic factor model. We can see in Figure 2.10 that the strategy was able to consistently generate a profit even when considering transaction costs.

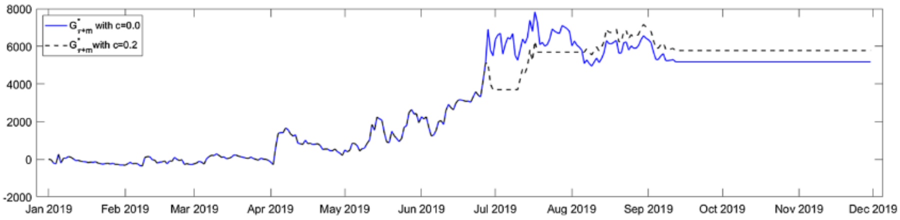


Figure 2.10: Net gains taking transaction fees into account [57]

Chapter 3

Project Plan

3.1 Fetch Historical Data

The first step of the project is to collect historical pricing data from UniSwap. As the UniSwap API does not provide this feature of historical prices, I plan to obtain this information by directly scraping every swap log out of each liquidity pool, then simply performing simple arithmetic to calculate the price with the *tick* attribute in the log.

3.2 Calculate Cointegrated Pairs

For each pair of exchange rate, calculate the sum of squared difference (SSD) and sort the pairs according to their SSD. Once the pairs have been sorted, calculate the cointegrated pairs and their parameters using the 2-step Eagle-Granger approach [59].

3.3 Implement the mean reversion strategy

Implement the simple mean reversion strategy as per mentioned in the background section. Using the cointegration factor calculated using the Eagle-Granger method as the hedge ratio, which remains constant.

3.4 Implement the backtesting system

The next step is to create the backtesting system that I will use to evaluate each strategy.

3.5 Apply Kalman Filter to update the hedge ratio

Building on the strategy above, instead of keeping the hedge ratio constant, the hedge ratio would constantly be updated as the prices change.

3.6 Apply ML Forecasting with the Kalman Filter to update the hedge ratio

The final strategy will use a combination of a machine learning learning forcasting algorithm to find opportunities that may not be discovered using the simple model. It is only after evaluating the strategy will we be able to justify whether this is a viable and profitable optimisation.

3.7 Evaluation

I then will evaluate each strategy and then hopefully optimise if required.

3.8 Timeline

30 Jan – Fetch Historical Data

6 Feb – Calculate Cointegrated Pairs

13 Feb – Implement the mean reversion strategy

27 Feb – Implement the backtesting system

6 Mar – Apply Kalman Filter to update the hedge ratio

27 Mar – Apply ML Forecasting with the Kalman Filter to update the hedge ratio

3 Apr – Evaluation, run strategy in real time

≈15 May – Final write up

Chapter 4

Evaluation Plan

4.1 Number Arbitrage opportunities found

One of the ways I will be evaluating the strategies I implemented will be finding the number of opportunities that each strategy finds as well as highlighting opportunities that resulted in both profits and loss. Furthermore, I will show how the P&L is effected by transaction/gas fees by plotting both the P&L including and excluding any fees.

4.2 Return on theoretical trading

One major problem with this project is the timeline, from background research, I can see that most papers evaluate their strategies over at least a 6 month period. As this limits development time and evaluation time, the majority of the evaluation will be run by a back-testing system and evaluating each transaction cost at each time for any transaction that I theoretically make to simulate the trade environment. Using this trading environment, I will run each strategy and plot it's revenue and return.

4.3 Return on actual trading

In addition to the trading on the in-sample data points, I plan to trade and evaluate the performance and returns of each strategy in real time.

4.4 Sharpe Ratio

Another metric I will use is the Sharpe Ratio, the reason for this is that it is important to evaluate how risky the strategies are compared to the risk free rate. The metric is given by the formula below:

$$\text{Sharpe Ratio} = \frac{R_p - R_f}{\sigma_p}$$

Where R_p is the return of a portfolio, R_f is the risk-free rate and σ_p is the standard deviation of the portfolio. I will calculate the Sharpe Ratio for both in-sample and out-of-sample results.

4.5 Performance of Signal Generation

I also plan to evaluate how long it takes for each strategy to generate a signal when prices change. In analysing this metric, we may be able to find times where once the signal is created, the opportunity could be lost hence by looking at this metric, we may be able to find processes that may decrease the signal creation time.

Chapter 5

Ethical Issues

5.1 Ethical Issues of Trading

5.2 Ethical Issues of Cryptocurrencies and Blockchain Technology

Bibliography

[1] "'I lost millions through cryptocurrency trading addiction'," *BBC News*, May 2021. [Online]. Available: https://www.bbc.com/news/uk-scotland-57268024

[2] C. Gondek, "What Big Companies Are Investing In Cryptocurrency?" [Online]. Available: https://originstamp.com/blog/what-big-companies-are-investing-in-cryptocurrency/

[3] "History of bitcoin exchanges and trading - Bit2Me Academy." [Online]. Available: https://academy.bit2me.com/en/history-exchanges-bitcoin-trading/

[4] "DEX to CEX Spot Trade Volume." [Online]. Available: https://www.theblock.co/data/decentralized-finance/dex-non-custodial/dex-to-cex-spot-trade-volume

[5] "Statistical Arbitrage: Definition, How It Works, and Example." [Online]. Available: https://www.investopedia.com/terms/s/statisticalarbitrage.asp

[6] M. Nofer, P. Gomber, O. Hinz, and D. Schiereck, "Blockchain," *Business & Information Systems Engineering*, vol. 59, no. 3, pp. 183–187, 2017.

[7] Y.-C. Liang, *Blockchain for Dynamic Spectrum Management*, 01 2020, pp. 121–146.

[8] "Components of Blockchain Network," Apr. 2021. [Online]. Available: https://www.geeksforgeeks.org/components-of-blockchain-network/

[9] "Merkle Tree in Blockchain: What is it and How does it work | Simplilearn." [Online]. Available: https://www.simplilearn.com/tutorials/blockchain-tutorial/merkle-tree-in-blockchain

[10] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," May 2009. [Online]. Available: http://www.bitcoin.org/bitcoin.pdf

[11] V. Buterin *et al.*, "A next-generation smart contract and decentralized application platform," *white paper*, vol. 3, no. 37, pp. 2–1, 2014.

[12] "Gas and fees." [Online]. Available: https://ethereum.org/en/developers/docs/gas/

[13] "Ethereum Virtual Machine (EVM)." [Online]. Available: https://ethereum.org/en/developers/docs/evm/

[14] "Introduction to smart contracts." [Online]. Available: https://ethereum.org/en/developers/docs/smart-contracts/

[15] "Smart Contracts - Industrial IoT Use Case Profile | IoT ONE Digital Transformation Advisors." [Online]. Available: https://www.iotone.com/usecase/smart-contracts/u88

[16] "What Is a DEX (Decentralized Exchange)? | Chainlink." [Online]. Available: https://chain.link/education-hub/what-is-decentralized-exchange-dex

[17] "Decentralized Exchange Platforms in Crypto Trading." [Online]. Available: https://www.gemini.com/cryptopedia/decentralized-exchange-crypto-dex

[18] "What is arbitrage?" Jul 2021. [Online]. Available: https://online.hbs.edu/blog/post/what-is-arbitrage

[19] G. Poitras, "Origins of arbitrage," *Financial History Review*, vol. 28, no. 1, p. 96–123, 2021.

[20] "Law of One Price: Definition, Example, Assumptions." [Online]. Available: https://www.investopedia.com/terms/l/law-one-price.asp

[21] P. Isard, "How far can we push the "law of one price"?" *The American Economic Review*, vol. 67, no. 5, pp. 942–948, 1977. [Online]. Available: http://www.jstor.org/stable/1828075

[22] J. Richardson, "Some empirical evidence on commodity arbitrage and the law of one price," *Journal of International Economics*, vol. 8, no. 2, pp. 341–351, 1978. [Online]. Available: https://www.sciencedirect.com/science/article/pii/0022199678900272

[23] I. Eyal, "The miner's dilemma," in *2015 IEEE Symposium on Security and Privacy*. IEEE, 2015, pp. 89–103.

[24] Z. Avarikioti, L. Heimbach, Y. Wang, and R. Wattenhofer, "Ride the lightning: The game theory of payment channels," in *International Conference on Financial Cryptography and Data Security*. Springer, 2020, pp. 264–283.

[25] G. Huberman, J. D. Leshno, and C. Moallemi, "Monopoly without a monopolist: An economic analysis of the bitcoin payment system," *The Review of Economic Studies*, vol. 88, no. 6, pp. 3011–3040, 2021.

[26] S. Athey, I. Parashkevov, V. Sarukkai, and J. Xia, "Bitcoin pricing, adoption, and usage: Theory and evidence," 2016.

[27] D. Easley, M. O'Hara, and S. Basu, "From mining to markets: The evolution of bitcoin transaction fees," *Journal of Financial Economics*, vol. 134, no. 1, pp. 91–109, 2019.

[28] C. R. Harvey, "Cryptofinance," *Available at SSRN 2438299*, 2016.

[29] E. Pagnotta and A. Buraschi, "An equilibrium valuation of bitcoin and decentralized network assets," *Available at SSRN 3142022*, 2018.

[30] J. Mo, "A Theoretical Model of Cross-market Arbitrage." [Online]. Available: https://w4.stern.nyu.edu/finance/docs/pdfs/PhD/mo-job-market.pdf

[31] N. Spiricheva, V. Baklykov, M. Tuhbadshin, and O. Mukhnovskiy, "Implementation of stock market arbitrage," in *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*, 2019, pp. 0055–0058.

[32] Q. F. Akram, D. Rime, and L. Sarno, "Arbitrage in the foreign exchange market: Turning on the microscope," *Journal of International Economics*, vol. 76, no. 2, pp. 237–253, 2008.

[33] Y. Aiba, N. Hatano, H. Takayasu, K. Marumo, and T. Shimizu, "Triangular arbitrage as an interaction among foreign exchange rates," *Physica A: Statistical Mechanics and its Applications*, vol. 310, no. 3-4, pp. 467–479, 2002.

[34] T. Ito, K. Yamada, M. Takayasu, and H. Takayasu, "Free lunch! arbitrage opportunities in the foreign exchange markets," National Bureau of Economic Research, Tech. Rep., 2012.

[35] I. Makarov and A. Schoar, "Trading and arbitrage in cryptocurrency markets," *Journal of financial economics*, vol. 135, no. 2, pp. 293–319, 2020.

[36] T. Crépellière, M. Pelster, and S. Zeisberger, "Arbitrage in the Market for Cryptocurrencies," Rochester, NY, Dec. 2022. [Online]. Available: https://papers.ssrn.com/abstract=3606053

[37] C. PAUNA, "Arbitrage trading systems for cryptocurrencies. design principles and server architecture," *Informatica economica*, vol. 22, no. 2/2018, pp. 35–42, 2018.

[38] N. Boonpeam, W. Werapun, and T. Karode, "The arbitrage system on decentralized exchanges," in *2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. IEEE, 2021, pp. 768–771.

[39] Y. Wang, Y. Chen, H. Wu, L. Zhou, S. Deng, and R. Wattenhofer, "Cyclic Arbitrage in Decentralized Exchanges," Jan. 2022, arXiv:2105.02784 [cs, q-fin]. [Online]. Available: http://arxiv.org/abs/2105.02784

[40] L. Heimbach, Y. Wang, and R. Wattenhofer, "Behavior of liquidity providers in decentralized exchanges," *arXiv preprint arXiv:2105.13822*, 2021.

[41] C. Krauss, "Statistical arbitrage pairs trading strategies: Review and outlook," *Journal of Economic Surveys*, vol. 31, no. 2, pp. 513–545, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/joes.12153

[42] H. Rad, R. K. Y. Low, and R. Faff, "The profitability of pairs trading strategies: distance, cointegration and copula methods," *Quantitative Finance*, vol. 16, no. 10, pp. 1541–1558, 2016. [Online]. Available: https://doi.org/10.1080/14697688.2016.1164337

[43] Y.-x. Lin, M. Michael, and G. Chandra, "Loss protection in pairs trading through minimum profit bounds: A cointegration approach," *Journal of Applied Mathematics and Decision Sciences*, vol. 2006, 08 2006.

[44] Z. Zhao, R. Zhou, and D. P. Palomar, "Optimal mean-reverting portfolio with leverage constraint for statistical arbitrage in finance," *IEEE transactions on signal processing*, vol. 67, no. 7, pp. 1681–1695, 2019.

[45] Q. Tran-Dinh, W. Michiels, and M. Diehl, "An inner convex approximation algorithm for bmi optimization and applications in control," *Proceedings of the IEEE Conference on Decision and Control*, 02 2012.

[46] Z. Zhao, R. Zhou, Z. Wang, and D. P. Palomar, "Optimal portfolio design for statistical arbitrage in finance," in *2018 IEEE Statistical Signal Processing Workshop (SSP)*, 2018, pp. 801–805.

[47] B. Alsadik, "Chapter 10 - kalman filter," in *Adjustment Models in 3D Geomatics and Computational Geophysics*, ser. Computational Geophysics, B. Alsadik, Ed. Elsevier, 2019, vol. 4, pp. 299–326. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128175880000106

[48] G. Vidyamurthy, *Pairs Trading: quantitative methods and analysis.* John Wiley & Sons, 2004, vol. 217.

[49] "How a Kalman filter works, in pictures." [Online]. Available: https://www.bzarg.com/p/how-a-kalman-filter-works-in-pictures/

[50] H. E. Dempsey, "Market Inefficiency: Pairs Trading with the Kalman Filter," Dec. 2017. [Online]. Available: http://www.hedempsey.com/papers/Pairs%20Trading%20with%20a%20Kalman%20Filter.pdf

[51] J. P. Nóbrega and A. L. I. Oliveira, "A combination forecasting model using machine learning and kalman filter for statistical arbitrage," in *2014 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2014, pp. 1294–1299.

[52] C. Krauss, X. A. Do, and N. Huck, "Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the s&p 500," *European Journal of Operational Research*, vol. 259, no. 2, pp. 689–702, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0377221716308657

[53] S. Moraes Sarmento, *A Machine Learning based Pairs Trading Investment Strategy*, 1st ed., ser. SpringerBriefs in Computational Intelligence. Cham: Springer International Publishing, 2021.

[54] T. G. Fischer, C. Krauss, and A. Deinert, "Statistical arbitrage in cryptocurrency markets," *Journal of Risk and Financial Management*, vol. 12, no. 1, 2019. [Online]. Available: https://www.mdpi.com/1911-8074/12/1/31

[55] D. Ju-Long, "Control problems of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016769118280025X

[56] R. Engle, "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, vol. 20, no. 3, pp. 339–350, 2002.

[57] G. Figá-Talamanca, S. Focardi, and M. Patacca, "Common dynamic factors for cryptocurrencies and multiple pair-trading statistical arbitrages," *Decisions in economics and finance*, vol. 44, no. 2, pp. 863–882, 2021.

[58] B. M. Blau, T. Griffith, and R. J. Whitby, "Comovement in the cryptocurrency market," *Economics Bulletin*, vol. 40, no. 1, p. 1, 2020.

[59] H. Rad, R. K. Y. Low, and R. Faff, "The profitability of pairs trading strategies: distance, cointegration and copula methods," *Quantitative Finance*, vol. 16, no. 10, pp. 1541–1558, 2016. [Online]. Available: https://doi.org/10.1080/14697688.2016.1164337