LUISS

Department of Economics and Finance
Equity markets and Alternative Investments

Performance Analysis of Statistical Arbitrage Portfolio

Advisor
Prof. Paolo Vitale

Co-advisor
Prof. Marco Morelli

716921
Niccolò Fabbri

Academic Year: 2020/2021

Contents

Chapter 1

Introduction

Harry Markowitz described diversification as *"the only free lunch in finance"*, risk can be reduced without sacrificing returns. However, empirical evidence suggests that there is another *"free lunch"*: arbitrages. Arbitrages occur when markets inefficiency generates an opportunity for risk-free profits.

The majority of trades nowadays are executed by algorithms. Institutional investors exploit arbitrages in seconds or even milliseconds, securing profits and restoring market efficiency.

This thesis will cover the topic of Statistical Arbitrage. This approach uses statistical techniques to infer relationships between assets and generate profits in their short-term deviation. However, such a strategy is not risk-free. It is exposed to model risk and security-specific risk. Relationships on which the model is based may be spurious or may not continue in the future due to changes in markets dynamics.

Our analysis wants to investigate if a Statistical Arbitrage trading strategy can take advantage of market inefficiencies to achieve consistent profits. In particular, we will develop an algorithm that backtests such a strategy to analyze its performance in the US stock market. Furthermore, alternative weighting schemes will be tested to assess the importance of asset allocation and risk budgeting in a market-neutral strategy.

Chapter 2 reviews the fundamentals of Modern Portfolio Theory, from the CAPM to APT, pointing out their features, with main benefits and drawbacks. Then, it covers the topic of long-short equity strategies and, therefore, market neutrality. Finally, it describes different approaches to pairs trading, focusing on quantitative ones, as tracking variance and cointegration.

Chapter 3 describes the entire trading strategy starting from the data polishing process. Subsequently, it will cover the topic of pair selection, introducing the concepts of stationarity, Engle-Granger test, and half-life. The last section of the chapter is dedicated to explaining how the algorithm opens or closes positions.

Chapter 4 examines the most popular weighting schemes lingering on their advantages and inefficiencies. Then, it describes alternative weighting schemes that may improve the performance of our strategy. In the end, it reviews the topic of risk budgeting, focusing on the equal risk contribution scheme and its implementation in our trading strategy.

The first section of chapter 5 explains the computation of the returns. Subsequently, it analyzes the performance of the strategy, comparing two different weighting schemes. In particular, the focus will be on returns, risk, and risk-adjusted measures.

Chapter 2

Overview of Pairs Trading

This chapter will examine pairs trading, from the economic models to the quantitative approach, to ensure a strong foundation and a deep understanding of the whole trading strategy covered in this thesis. First, we will start from the most known asset pricing models and then analyze a long-short strategy with its advantages and drawbacks. In the end, we will review the various approaches of pairs trading, lingering on the statistical concepts behind them, helpful to grasp the whole dissertation.

2.1 Asset pricing models

2.1.1 Capital Asset Pricing Model

The CAPM is the best-known asset pricing model which describes the relationship between systematic risk and expected return on assets. With the formula:

$$E[R_i] = R_f + B_i(E[R_m] - R_f) \tag{2.1}$$

where $E[R_i]$ is the expected return of the asset, R_f is the risk-free rate, and B_i is the beta of the stock, representing the risk exposure of the i asset. Market beta is 1; if the i-asset beta is greater than 1, it is considered riskier than the market, and it will increase the overall risk of the portfolio. On the contrary, if the i-asset has a beta smaller than one is considered low-risk; the smaller the β, the closer the expected return gets to the risk-free rate.

CAPM finds its use in asset allocation, given an input of expected return and a covariance matrix, the mean-variance optimizer will return the efficient frontier, representing the set of portfolios with the best risk-return trade-off. We can obtain the efficient risky portfolio in terms of Sharpe-Ratio by drawing the tangent capital allocation line to the frontier. For example, in the Fig 2.1 below, we can see the efficient frontier considering FAANG [1] companies and S&P500 as a proxy of the market considering the period 2018-2020[2], with the weights respectively in table 2.1.

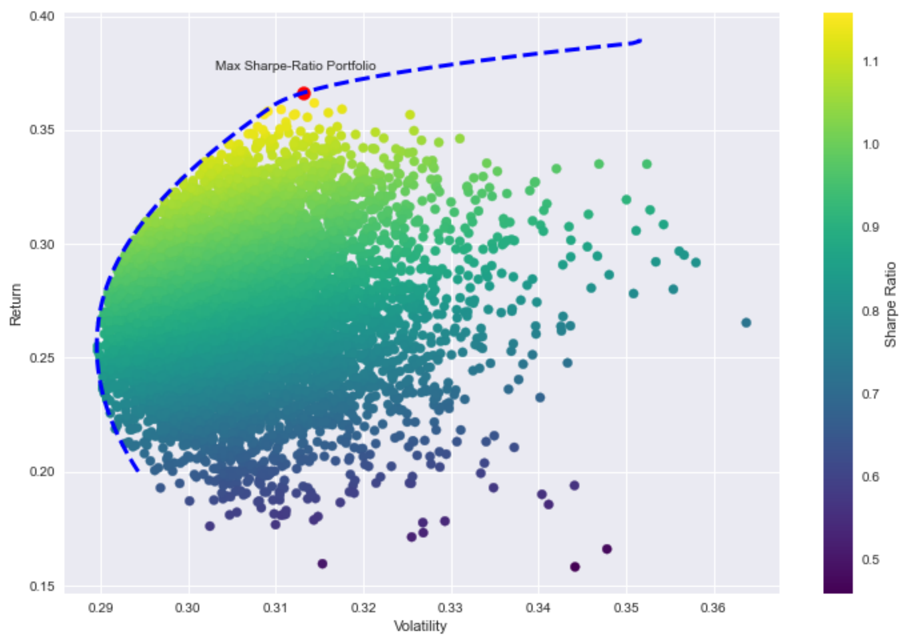

Figure 2.1: Efficient Frontier and Max Sharpe-Ratio portfolio of FAANG Companies

ticker	AAPL	AMZN	FB	GOOG	NFLX
return	0.38	0.32	0.06	0.12	0.26
weights	58%	37%	$\sim 0\%$	$\sim 0\%$	4%

Table 2.1: Returns and optimal weights in the period 2018-2020

The pure mathematical resolution of the maximization problem will lead to unreasonable weights due to lack of robustness of Markowitz's model. The

[1]Facebook, Amazon, Apple, Netflix, and Google
[2]Yahoo finance data

ideal allocation in our case would be to invest 95% of the budget in just two stocks out of five (AAPL and AMZN), this will obviously deface the benefit of diversification. Such allocation is sample-biased and in a forward-looking perspective, worthless. Professionals tend to avoid the Maximum Sharpe-Ratio portfolio because of the input sensitivity of the model: an error on 1% of an expected return estimation can strongly upset the asset allocation. Instead, asset managers focus more on Global Minimum Variance (GMV) portfolios that do not require expected returns estimates. Therefore, the weights depend only on risk parameters; it is well known that it is much easier to forecast volatility than returns.

There are several assumptions behind the CAPM: investors are rational, all assets are publicly held and trade on public exchanges, investors can borrow or lend at risk-free rates, and take short positions on assets. There are no taxes or transactions costs. This model despite its empirical shortcomings remains widely used[3].

2.1.2 Arbitrage Pricing Theory

Arbitrage Pricing Theory was initially introduced in 1976 by Stephen A. Ross. The model assumes that security mispricings cannot lead to a risk-free profit; there are no arbitrages: if an arbitrage opportunity opens, there would be strong pressure on prices that would close the opportunity immediately. APT is a multi-factor model in which risk premiums derive from risk exposure to multiple risk sources [4]. The key idea behind the APT is that securities with the same sensitivity to macroeconomic factors should have the same expected return. The formula that describes the model is:

$$R_i = R_f + \beta_1 R_1 + \beta_2 R_2 + ... + \beta_n R_n + \epsilon = R_f + \sum_{k=1}^{n} \beta_k R_k + \epsilon \qquad (2.2)$$

where R_i is the return of the i-asset,R_f is the risk-free rate, β_k is the sensitivity of the i-asset to the k-factor, R_k denote the return contribution of each risk factor and ϵ represent the idiosyncratic risk on the stock that is not explicable by the model.

The APT, likewise CAPM, is used in capital budgeting, security valuation, and investment performance evaluation. The former does not rely on

[3]Bodie et al., 2009, chap 9
[4]Bodie et al., 2009, chap 10

the market portfolio with all assets that, in reality, is not observable. Instead, it works with a diversified index portfolio like S&P500. The latter to be robust requires that each investor invests, maximizing the Sharpe-Ratio, the APT instead, relying on the fact that only a few arbitrageurs can bring the market to equilibrium and rule out arbitrage opportunities. The concepts of arbitrage due to mispricing and the APT will be helpful in the following sections to explain the purpose of the trading strategy covered in this thesis.

2.2 Long-Short strategies

Long-Short Equity finds its roots in Alfred Winslow Jones, the first hedge fund manager. These strategies consist of combining long positions in underpriced stocks and short positions in overpriced stocks. Ideally, the investor would benefit from both positions and will also minimize his market exposure. Although long-short equity funds can profit even from a declining market, the manager can also be partly wrong, at the end, what matters is that the net profit is positive. For example, let us consider this scenario: we have two stocks A and B, at t_0 the prices are 50$ and 75$ respectively. According to our forecasts, we think A is overvalued, and B is undervalued; therefore, we open a long/short position.

- Long 75$ B
- Short 50$ A
- Net exposure : 25 $

At t_1, the price of A will be 30$, and the price of B will be 70$, stock A depreciated and generated a profit. On the contrary, stock B generated a loss.

	Rate	**PnL**
Stock A	+40%	+20$
Stock B	-6.66%	-5$
Overall	+16% - 4%	15$

Table 2.2: Profit and Loss of Stocks A and B

The total profit is 15$. As a portion of the risky capital which is 125$ (considering the gross exposure), we will obtain a total net return of 12%

even if our forecast on Stock B turned out to be wrong. In addition if stock A and stock B were positively correlated, the overall risk of the position would be lower than a long-only position in these two stocks.

The main goal of these strategies is to gain profits and at the same time reducing market risk; they aim to be *beta neutral* in a CAPM perspective. For example, if we consider two stocks of the last example, Stock C with $\beta_C = 1.3$ and Stock D with $\beta_D = 1.2$, and we assume that the stocks have the same price[5]. Then, since the beta of a portfolio is the weighted average of the betas of its components, the market exposure of the long-short portfolio with a long position in C and a short position in D is:

$$\beta_{L/S} = \frac{1}{2}\beta_C - \frac{1}{2}\beta_D \tag{2.3}$$

$$\beta_{L/S} = \frac{1}{2}(1.3) - \frac{1}{2}(1.2) = 0.05$$

Long-Short positions are usually constructed to obtain a negligible beta and therefore minimize the exposure to the market, leading to *market neutral* portfolios. Market-neutral strategies aim to avoid market risk, and in practice, their returns should be uncorrelated with market returns. Although these strategies do not necessarily imply that portfolios are *dollar neutral*, they will cost less than a long-only portfolio since you short sell, but the positions do not need to match in order to obtain a 0$ investment.

However, Long-Short equity strategies are not the panacea of equity investing; they have lots of disadvantages too[6]:

- Higher trading costs due to short selling.
- Lag in bull markets, since the strategy involves short selling, will earn less than a long-only strategy in a bullish market, although that is the price the investor pays to reduce his market exposure.
- Short selling is usually allowed at a price higher than the "last" transaction price.

There are several approaches that fund managers in long-short equity funds can use. Many funds are sector-oriented; they operate just in a specific

[5]to make calculations easier
[6]Lhabitant, 2007, chap 7

sector: healthcare, technology, real estate, and energy are the most common. This behavior is justified by the expertise of the manager in that particular field.

The most famous approach in long-short equity is the valuation-based approach; it uses fundamental analysis to understand whether a stock is overvalued or undervalued. The idea behind this approach is the intrinsic value. In the medium term, the price has to converge to its intrinsic value and meet the manager's expectation. In order to determine the intrinsic value, most analysts use DCF models. They forecast future cash flows of the company, discount them with a specific return rate, and obtain the Present Value. This process is rich in assumptions, and on account of that, the models output a range of values, not a unique one. Analysts, when deriving intrinsic value, tend to use relative valuation metrics such as P/B (price to book value), P/E (price to earnings), and M/B (book to market value). These metrics are applied to comparable companies to assess whether there is an underpricing or overpricing. Fundamental analysis also uses balance sheets and income statements to verify the overall health of the company. It also uses qualitative measures that refer to the strategic positioning of the company, its management, the industry, and the business model[7]. Since these forecasts have a certain degree of subjectivity, the managers set a margin of safety in their exit and entry points. When a stock enters the portfolio, it is constantly monitored and assigned a target exit price, stop loss, and holding period.

As Lhabitant (2007) says, this valuation-driven approach is in total conflict with the EMH (Efficient Market Hypothesis); indeed, EMH states that all the possible information is available and reflected on market prices. Therefore, hedge fund analysts would not gain any value added by performing fundamental analysis.

Another Long-Short equity approach that is catching on is the quantitative approach. The quantitative analysis aims to identify the most tradable among a broad set of securities concerning the trading strategy. Quant models rely on mathematical and statistical techniques for asset allocation, stock picking, risk management, and evaluating potential trades. The following section will focus on a particular long-short equity strategy whose name is Pairs Trading and the quantitative techniques that rule this strategy.

[7] Lhabitant, 2007, chap 7, p. 171

2.3 Pairs Trading approaches

The first practice of statistical pairs trading is attributed to Wall Street quant Nunzio Tartaglia who was working in Morgan Stanley in the mid-1980s[8]. He introduced a trading technique that aims to identify pairs of securities whose prices move together, and if the spread between them widens, an arbitrage opens. However, there is no certainty that stock prices preserve their past behavior in the future; therefore, pairs trading may be unprofitable.

2.3.1 Naive distance approach

We will review the approach used in Gatev et al. (2006) by far the most famous paper on Pairs Trading. One of the simplest quantitative methods uses a distance measure; this section will focus on the tracking variance: the estimated average price distance, the sum of squared deviations of normalized prices divided by the number of periods. For example, given two stocks A and B, the tracking variance in the period going from 1 to T is defined as:

$$TV = \frac{1}{T}\sum_{t=1}^{T}(Q_t^A - Q_t^B)^2 \tag{2.4}$$

Where Q_A and Q_B represent the normalized prices, $Q_t = P_t/P_1$ of the two stocks and t represents the time. We select pairs with this measure, choosing those for which such measure is the smallest. Once the pairs are identified likewise Gatev et al. (2006) we can use the following strategy:

1. Track the spread Δ between normalized prices; this will create a time series of the spread.

2. Identify the threshold, in our case $2\sigma_\Delta$ where σ represents the standard deviation.

3. When the threshold is triggered, sell the most expensive stock and buy the cheapest.

4. Close the position when the stock prices cross (spread goes back to 0) or when the prices widen and stop-loss is triggered.

[8]Vidyamurthy, 2004, chap 5 p.74

Let us propose an illustrative example; we want to implement the strategy for the first ten stocks in S&P500 sorted by market capitalization: *Apple, Microsoft, Amazon, Facebook, Google, Tesla, Visa, NVIDIA Corporation, JPMorgan Chase, Johnson Johnson*. We have downloaded historical prices from yahoo finance for the period 2018-2019(included). First, we will normalize our dataset by dividing itself for the first row, and then with a simple function in python, we can compute the tracking variance for each pair in our basket of 10 stocks. Once the code is executed, we can see that the pair GOOGL-JPM has the lowest tracking variance, particularly 0.00805.[9] Figure 2.2 represents GOOGL and JPM normalized prices in the same plot, it's not unreasonable to say that these stocks move together.

date	GOOGl	JPM	P_G	P_J
2018-07-13	Sell	Buy	1204.42	96.99
2018-09-19	Buy	Sell	1174.27	107.26
2019-03-12	Sell	Buy	1197.25	96.31
2019-05-01	Buy	Sell	1173.31	107.42

Table 2.3: Trades executed with dates and prices

In Fig. 2.3 we can see the spread (GOOGL-JPM), if at time t the absolute value of the spread takes values greater than $2\sigma_\Delta$ (0.01148 in our case) the threshold is triggered. In our example, this happens two times, on 2018-07-13 and 2019-03-12. On both days, GOOGL is more expensive than JPM (in a normalized scenario), which means that we have to short sell Google and buy JPMorgan. Positions are closed on 2018-09-19 and 2019-05-01, respectively. Figure 2.4 depicts normalized prices and the positions taken in the two securities (see trades details in table 2.3). Trades stay open for 47 and 35 days for a total of 82 out of 503 trading days in our sample, which roughly represents 16% percent of the total time. This may be due to the smallness of our sample that influenced the quality of the pairs and the co-movement relationship. The "most tradable pair" in our dataset could results in the average in greater samples.

Moreover, the thresholds are not designed for this specific pair, and they

[9]Given the explanatory purpose of this example, the strategy will be tested in-sample. However, to validate this methodology, we should separate the estimation window, which is used to identify the pair, and the performance window, which is used to analyze the performance of the strategy

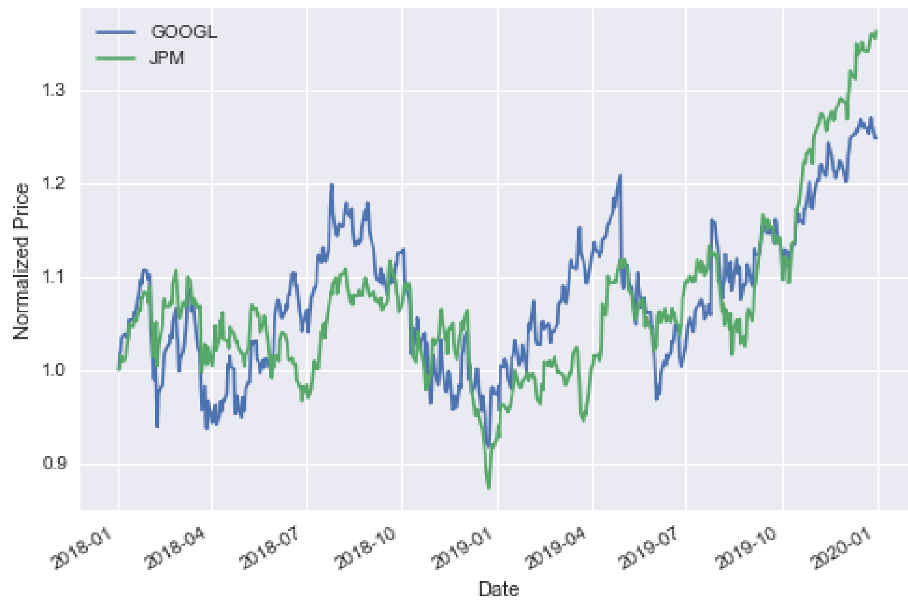

Figure 2.2: Normalized prices of Google and JPMorgan Chase & Co. stocks,2018-2019

could be too "strict" for the statistical properties of these time series. However, it is not recommended to optimize each pair's thresholds to avoid overfitting and sample dependency. In the end, even if just two trades were executed, our in-sample analysis turned out to be profitable: all four open positions have been close in profit.

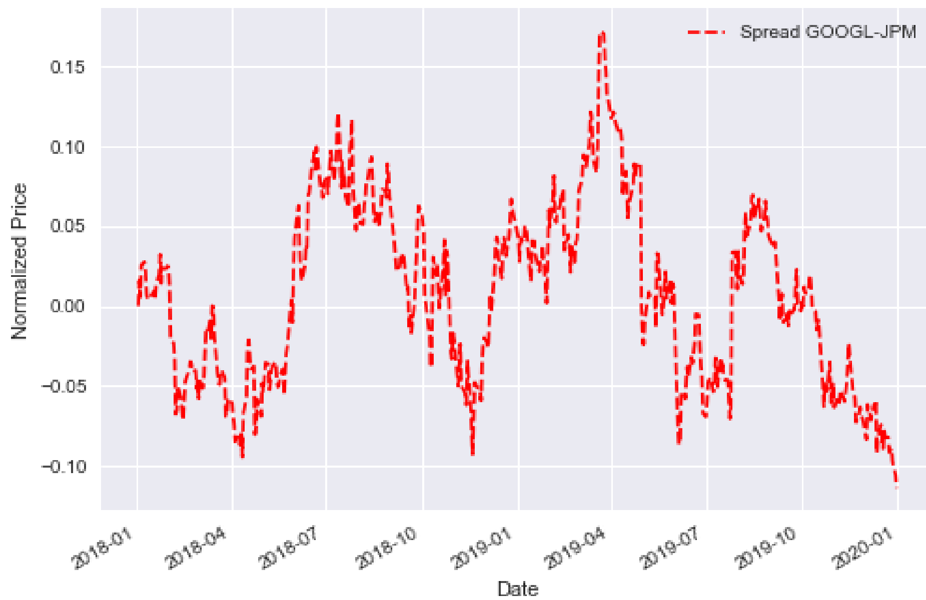

Figure 2.3: Normalized spread of Google and JPMorgan Chase & Co. stocks, 2018-2019

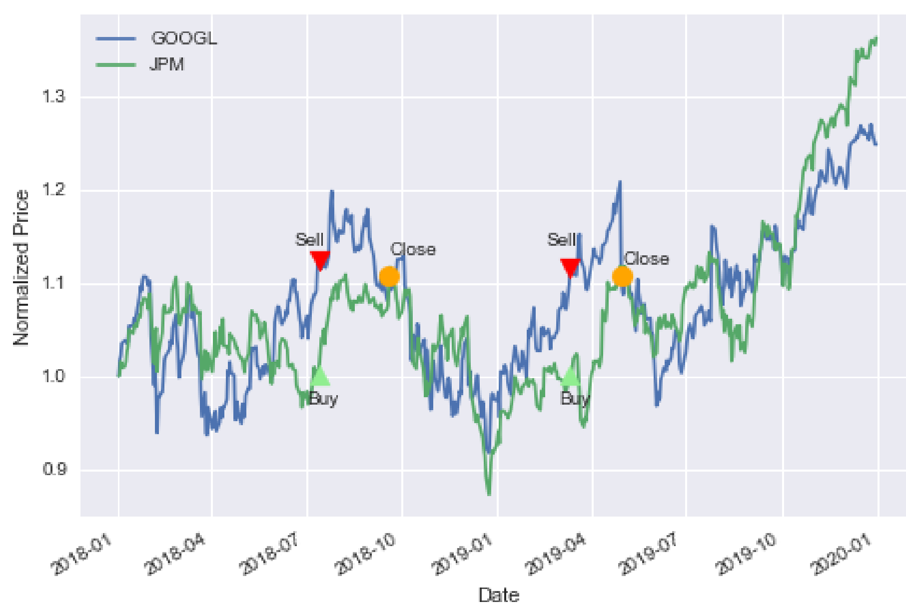

Figure 2.4: Normalized prices and trades executed

2.3.2 Cointegration approach

A time series is stationary if the stochastic process generating the series has time-invariant parameters. Stationary times series are easy to forecast, and they are mean-reverting, see Fig. 2.5. Clearly, if stock prices were mean-reverting, we could buy when the price is below the mean and, on the contrary, sell the security when it is above it. This would guarantee us a risk-free profit[10], relying on the fact that the price will converge to mean. Empirically we can model a generic stock price time series with a random walk, see Fig. 2.5, which is the simplest example of an integrated process $I(1)$. Random walks sequences are unpredictable. For this reason, the work of Engle and Granger (1987) turned out to be one of the most valuable discoveries both in theory and in practice since many funds adopted a trading strategy based on their model. Statistical Arbitrage based on cointegration analysis relies on the non-stationarity of stock prices. In particular, time series must be an integrated process of order 1 $I(1)$. An example of such process can be seen in figure 2.5 and 2.6.

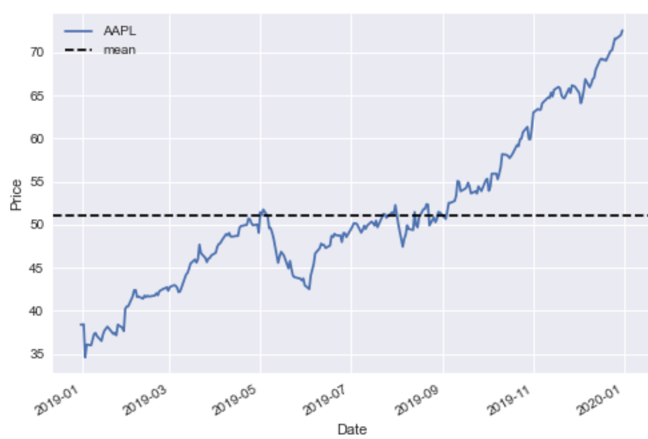

Figure 2.5: AAPL price 2019, non-stationary process, yahoofinance

[10]hence the name statistical "arbitrage"

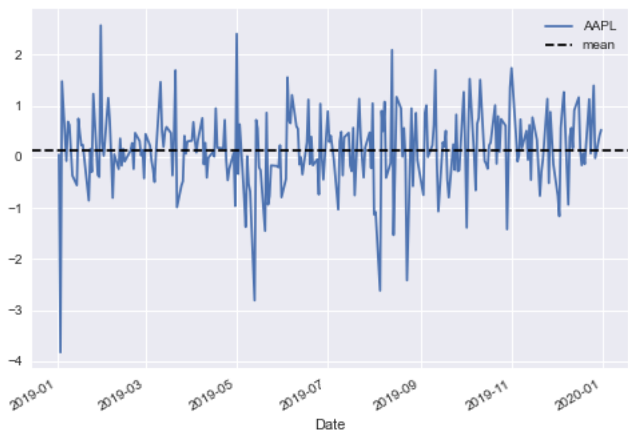

Figure 2.6: First differences AAPL price 2019, stationary process, yahoofinance

This means that the process itself it's not stationary but its first difference is. Two processes $I(1)$ are cointegrated if there's a linear combination of them that is stationary $I(0)$, if two time series are cointegrated they comove. In brief, if y_t and x_t are historical stock prices there are two coefficients α and β such that:

$$\alpha x_t + \beta y_t = z_t \text{ where } z_t \text{ is } I(0) \tag{2.5}$$

A long-term equilibrium relationship links these processes. Therefore, the investor should take advantage of the short-term deviations from the equilibrium. The goal is to find cointegrated securities where the spread z_t has a high degree of mean reversion, estimate the cointegration coefficient, and profit on temporary mispricings. Cointegration and its implementation in a trading strategy will be explored in more detail in future chapters since the purpose of this section is to provide the reader with all the necessary tools needed for a proper understanding of the entire thesis.

Chapter 3

Trading Strategy

3.1 Data preparation

The data analysis and the backtesting of the strategy will be performed on Jupyter Notebook with Python. We will use daily data of stock prices of all the components of S&P500 in the period 2014-2021Q2, provided by Bloomberg. Since many companies in the index were included in it during our analysis period, to avoid biases, we will remove every stock that it's been added or removed in that period. This cleaning process will lead us to a DataFrame with 476 stocks rather than 500. To validate the strategy in a forward-looking perspective, we will split the dataset into two parts to analyze the in-sample and the out-of-sample performance. The validation will be performed on a rolling basis. The estimation window will be five years long and the out-of-sample three months. Every quarter the in-sample dataset will slide, taking into account information of the out-of-sample dataset, and the parameter estimates will be updated in a rolling way.

3.2 Pairs selection

In order to create an active portfolio, we need to select several pairs; the number can range from 15-30 to maximize diversification-transaction cost trade-off. As mentioned in section 2.3.2 our criterion to identify pairs is Cointegration. First of all, since Cointegration assumes that the processes are integrated, we have to remove from our basket of securities the stocks which prices form a stationary process, see fig. 2.6. Obviously, to test the stationarity of a time series, the plot is not enough. We need to run a unit root test and check whether its parameters change over time.

3.2.1 Unit root tests

There are several unit root tests; in our analysis, we will use the Augmented-Dickey-Fuller test. To understand why unit-roots are linked to stationarity, let us suppose to have an AR(1) model described by the formula:

$$y_t = \phi y_{t-1} + \varepsilon_t = y_0^t + \sum_{k}^{t-1} \phi^k \varepsilon_{t-k} \tag{3.1}$$

Where ϕ is the coefficient that scales the lagged value at $t-1$, ε_t is the error term at time t and the last part of the equation is the MA representation that equals the AR(1) model[1]. The expected value and the variance are:

$$E(y_t) = \phi^t y_0 \tag{3.2}$$

$$Var(y_t) = \sigma^2(\phi^0 + \phi^2 + \phi^4 + ... + \phi^{2(t-1)}) \tag{3.3}$$

The variance of each ε_t is assumed to be the same σ^2, so we can factor it out. How do these two parameters change when ϕ does? We have three scenarios, $|\phi|$ can be greater, smaller, or equals to 1. In the first case, the time series "explodes" over time and will not be stationary. If the absolute value of the scaling coefficient is less than one, the expected value in the eq. 3.2, will converge to 0. Accordingly, the variance in eq. 3.3, represents a geometric series that converges to $\frac{\sigma^2}{1-\phi^2}$. In this case, both the expected value and the variance do not depend on t, they are constant over time, and therefore this condition leads to stationarity. The last case, where ϕ equals 1 or -1, is named unit root. Even if the expected value, in our simple AR(1) scenario, would result time-invariant, the variance would be $t\sigma^2$ and, therefore, nonstationary.

Our analysis will test the null hypothesis of having unit-roots against H_1 that states there are no unit-roots. The ADF test assumes that the times series is a generic AR(p) model:

$$y_t = \mu + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t \tag{3.4}$$

[1]this representation will make more clear the math behind unit roots

We subtract y_{t-1} from both sides

$$\Delta y_t = \mu + \delta y_{t-1} + \sum_{i=1}^{p} \beta_i \Delta y_{t-i} + \varepsilon_t \tag{3.5}$$

Test hypotesis will be:

- $H_0 : \delta = 0$
- $H_1 : \delta < 0$

In order to perform this statistic test, we need to compute t-statistic $t_{\hat{\delta}}$ of the δ and then compare it with the Dickey-Fuller distribution. If $t_{\hat{\delta}}$ is greater than the DF critical value, we cannot reject the null hypothesis; otherwise, we will reject it and conclude that the time series is stationary. We will run the ADF test on the initial in-sample DataFrame to identify the suitable time series for the analysis. The first period used for parameters estimation goes from January 2014 to December 2018. We find out that 13 out of 476 time-series are stationary, so we need to drop them. This process will lead to a cleaned dataset of 463 stocks.

In the table 3.1 below, we can see the companies whose stock prices are generated by a stationary process sorted by their market capitalization[2]in billions. The sum of the capitalizations is 579.989 billions. Since the total capitalization of the S&P500 is around 37 trillions [3], we removed approximately 1.5% of the entire index. In addition, we can see that the companies' betas range from 0.3 to 1.7, underlying that the stationarity does not depend on the company's riskiness.

[2]Data provided by Yahoo-finance!
[3]accordingly to 2021Q3

Company	Market Cap	Beta
SEMPRA	4.767	1.39
JUNIPER NETWORKS	9.550	0.83
AMERICAN AIRLINES GROUP	12.749	1.79
OMNICOM GROUP	15.939	0.94
GENUINE PARTS	17.542	1.14
BOSTON PROPERTIES	17.963	1.24
KANSAS CITY SOUTHERN	26.918	1.08
PPG INDUSTRIES	37.625	1.13
METLIFE	52.836	1.31
DOMINION ENERGY	63.782	0.36
COLGATE-PALM.	66.233	0.60
GENERAL MOTORS	71.062	1.33
WELLS FARGO & CO	183.023	1.36
Total/Range	579.989	0.36-1.79

Table 3.1: Companies removed from the analysis sorted by Market Capitalization

3.2.2 Cointegration testing and spread series

To find the pairs in our sample, we need to run the Engle-Granger Test on the whole DataFrame. The E-G tests for no-cointegration of a univariate equation and is a two-step process:

1. Determination of the linear relationship.
2. Stationarity testing on the spread series (with ADF test).

The null hypothesis is no cointegration. If the p-value is small, below a critical size, we can reject the hypothesis that there is no cointegrating relationship. Given 463 different stocks, we have 106953 possible pairs. Once we have iterated the statistical test over every possible pair of time series, we find 5306 pairs at a confidence level of 5%. Therefore for each pair, we can construct the spread series z_t:

$$y_t - \beta x_t = z_t \text{ where } z_t \text{ is stationary} \tag{3.6}$$

Where y_t and x_t represent two time-series, and β represents the cointegration coefficient. To compute the latter, we will use OLS, performing a regression of y_t against x_t. The beta of the regression is the quantity of x necessary to hedge 1 unit of y and therefore to maintain the stationarity relationship. Let us propose an illustrative example: we take the pair consisting of *Expeditors International of Washington Inc* and *CDW Corporation*. We can construct the spread series in three steps:

1. run two regressions swapping the dependent and independent variable and find two coefficients.

2. pick the larger, from a numerical viewpoint, reduce precision errors, and designate the stock with lower volatility as the independent variable[4].

3. subtract the independent variable multiplied by beta from the dependent one.

In our case, the cointegration coefficient is 1.86, with *DWS Corporation* as the dependent variable y_t and *Expeditors International of Washington Inc.* as the independent one, x_t. The figure 3.1 depicts the spread series for the first in-sample period going from 2014-2019.

This vast number of pairs suggests the presence of data snooping bias: the risk of having false positives dramatically increases with iteration. To remove this bias and select an ideal number of pairs, we will use the half-life criterion, which will allow us to choose the best suitable pairs for the strategy. Furthermore, since we need to find the most tradable ones, relying on the fact that we trade on a stationary process, we want that process to revert to its mean as fast as possible to execute more trades. Therefore, we will measure the tradability of a pair by the half-life of its spread series. The half-life is defined as the number of periods required for the impulse response to a unit shock to a time series to dissipate by half. It is widely used to quantify the degree of mean reversion. In AR(1) models and for stationary data, the half-life is described as:

$$h = -\frac{log(2)}{log(\rho_1)} \tag{3.7}$$

where ρ_1 is the autocorrelation of the spread series at lag 1. Since we use daily data, the half-life is expressed in days. Ideally, we want half-life to be as small as possible to maximize the number of trades. Therefore, we sort all the

[4]Vidyamurthy, 2004, chap. 7 p.108

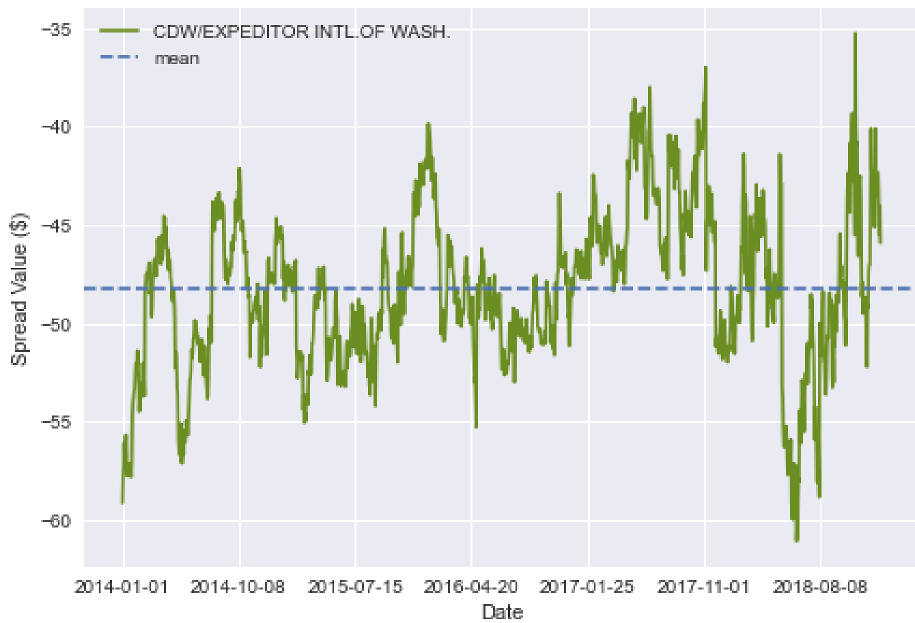

Figure 3.1: Spread Series of DWS Corporation and Expeditors International of Washington Inc.

5306 pairs by their half-life measured on the in-sample dataset and assume that they will keep their mean-reverting behavior in the future. Thus, our analysis will select the top 20 pairs, in which the spread will revert back to its mean value in fewer days than others, the most tradable ones according to this criterion. In table 3.2 we can see the details of each pair with the starting beta[5]. We can infer that series with the properties cited above are likely to remain the same, even in different pairs. For example, *Expeditors International of Washington Inc.* is present in four different pairs, *Crown Castle* in three and *CDW Corporation*, *Mastercard*, and other companies in two. The first one is a global shipment company. Crown Castle is a real estate investment trust and provider of communications infrastructure in the United States. CDW Corporation is a provider of technology products and services for businesses, and Mastercard is a financial-tech company. Their market capitalizations are, in billions: 20, 81, 26, and 338. We can see that their market cap varies a lot; we are not tilted towards small-cap o large-cap

[5]the cointegration coefficient regarding the first in-sample estimate.

stocks. However, apart from *Expeditors International of Washington Inc.*, we have selected companies related to the tech industry. For this reason, our portfolio will be exposed more to it.

Y-series	**X-series**	**Beta**
CDW	EXPEDITOR INTL.OF WASH.	1.86
XYLEM	AFLAC	2.51
CHAS.RVR.LABS.INTL.	CROWN CASTLE INTL.	1.63
CINTAS	EXPEDITOR INTL.OF WASH.	4.20
SERVICENOW	MASTERCARD	0.99
ADOBE (NAS)	MASTERCARD	1.48
INTUITIVE SURGICAL	THERMO FISHER SCIENTIFIC	3.27
SVB FINANCIAL GROUP	TEXTRON	6.75
BECTON DICKINSON	CROWN CASTLE INTL.	3.27
UNITEDHEALTH GROUP	CDW	3.09
THERMO FISHER SCIENTIFIC	NEXTERA ENERGY	5.49
STRYKER	INTERCONTINENTAL EX.	2.46
WATERS	WILLIS TOWERS WATSON	1.97
CHARTER COMMS.CL.A	COMCAST A	14.91
AMER.ELEC.PWR.	EVERSOURCE ENERGY	1.20
CITRIX SYS.	EXPEDITOR INTL.OF WASH.	1.92
MSCI	EXPEDITOR INTL.OF WASH.	3.90
WATERS	CROWN CASTLE INTL.	2.61
L3HARRIS TECHNOLOGIES	MAXIM INTEGRATED PRDS.	3.09
DANAHER	INTERCONTINENTAL EX.	1.21

Table 3.2: Selected pairs for the analysis

3.3 Trading signals

In this section, we will cover the topic of trading signals. We will explain how they are generated and their functioning. To backtest the trading strategy, we need a specific rule that, when is triggered, tells us to open or close a position. In our case, we are aware that the spread series is stationary, with a high degree of mean reversion; indeed, we will base our signals on the latter. In particular, we will standardize every spread series with the equation 3.8

to create a level playing field and set up a common rule.

$$ns_t = \frac{s_t - \mu_s}{\sigma_s} \tag{3.8}$$

Where ns_t represent the normalized spread at time t, s_t is the value of the spread at time t, μ_s represent the mean of the spread series and σ_s is its standard deviation. Once the spread series is normalized, its mean will be 0. Since we want to take advantage of the temporary mispricings, we will trade the assets when the spread is above or below a certain threshold because we rely on the fact that it will go back to 0. We want to avoid overfitting, and for this reason, we will set a unique threshold of 1. When the absolute value of the spread series takes a value greater than one, a position on the spread will be opened. Opening a position on the spread means opening a long and a short position simultaneously. Specifically:

- long spread : buy the *Y-asset* and short β on the *X-asset*.
- short spread : short the *Y-asset* and buy β on the *X-asset*.

In particular:

- If the spread takes a value lower than -1, we will open a long position on the spread.
- If the spread takes a value greater than 1, we will short the spread.
- The exit point will be 0.

For example, if the spread's value is greater than 1, we will short the spread; when the spread crosses 0, we will long the spread to close the position. Furthermore, we can open/close just one position per day: if we have a short position and the spread rockets down to -1, first we will close the position, and then, the following day, we will open a long position on the spread. We designed a function that creates $0, 1, -1$ values when the threshold is triggered on a corresponding day. If on *X-day* a long position is opened, writes 1; if a short position is opened -1, when none of the conditions are met, writes 0. Thus we can manage open positions. If on *X-day* my cumulative sum is 1, I know a long position is opened, and to close it, I need to go short. On the other hand, if my cumulative sum is 0, I do not have open positions.

At the beginning of the chapter, we mentioned that the validation of our strategy would be done on a rolling basis. In particular, the strategy will take into account information of the out sample period every three months; pairs will remain the same in every period. However, the cointegration coefficient will be updated every quarter. In our case, the first in-sample dataset goes from 2014Q1 to 2018Q4, and the out of sample will be 2019Q1. Once tested the performance over that period, the in-sample DataFrame will slide, thus becoming 2014Q2-2019Q1, as shown in the table 3.3.

out-of-sample	2019Q1	2019Q2	...	2021Q2
in-sample	2014Q1-2018Q4	2014Q2-2019Q1	...	2016Q2-2021Q1

Table 3.3: rolling in-sample and out-of-sample DataFrames

Since we update quarterly the cointegration coefficients, which represent the quantity of *X-asset* to buy or sell, if, at the end of a quarter, we have open positions, we will close them to avoid position mismatching.

To summarize, let us propose an example: consider the pair consisting of *Xylem* and *Aflac*. This pair's half-life has been estimated at 14 days. We want to generate trading signals regarding the first out-of-sample DataFrame(2019Q1). First of all, we compute the cointegration coefficient and the spread series. Secondly, we standardize the spread series with the formula 3.8, then we check whether the thresholds are triggered, and we open positions accordingly. The spread series hits the threshold five times, two times triggering a long signal and three times triggering a short one. Thus, four trades out of five are closed automatically, as can be seen in Fig. 3.2. Trades details are in table 3.4.

The process described in the last example will be done for each pair and the entire out-of-sample period. This chapter has explained in depth the trading strategy with its strengths and its limits. We want to maximize the performance of our trading strategy. For this reason, in the next chapter, we will examine alternative weighting schemes that may or may not enhance the return or risk profile of the strategy. Finally, we will compare a naive weighting scheme vis a vis risk-based indexation. Since returns depend on the weighting scheme of the portfolio, for this reason, their computation will be covered in the following chapters.

Figure 3.2: Xylem-Aflac spread series with signals

Day	P_{XYLEM}	P_{AFLAC}	Signal	Position
3	66.94	44.92	-1	short
11	68.97	46.81	1	close
15	69.27	47.39	1	long
20	70.49	47.11	-1	close
23	70.15	47.94	1	long
38	74.80	48.75	-1	close
49	76.89	49.07	-1	short
52	76.50	49.70	1	close
56	78.18	49.46	-1	short

Table 3.4: Trading details XYLEM-AFLAC pair regarding 2019Q1

Chapter 4

Weighting Schemes

Since our strategy relies on trading 20 different pairs, the topic of budget allocation is crucial. We desire to allocate the budget to the best-performing assets, both in terms of risk and return. This chapter will analyze the most popular weighting schemes, with their advantages and disadvantages. In primis, the most traditional allocation techniques will be treated, describing their success and its causes. Then, the chapter will cover risk budgeting with a particular focus on the risk parity portfolio. Ultimately, we will cover the implementation of the equal risk contribution principle in our strategy.

4.1 Tradional weighting schemes

In today's day and age, most funds base their portfolio optimizations on Modern Portfolio Theory. In his 1952 paper, Markowitz described portfolio selection with a simple rule: "The investor does (or should) consider expected return a desirable thing and variance of return an undesirable thing.". He stated that there is no optimal portfolio but a set of efficient portfolios. Tobin in 1958 showed that one portfolio dominates among the other in the presence of the risk-free asset: the tangency portfolio. The latter is the portfolio that maximizes the Sharpe Ratio. Later, Sharpe developed the CAPM, which we described in Chapter 2. He discovered that the prices of the assets are such that the tangency portfolio is the market portfolio: the portfolio containing every asset weighted by market capitalization. Thus, in an efficient scenario, the investor would hold the market portfolio. Combined with EMH developed by FAMA in 1965, these theories led to the thought that an investor cannot beat the market consistently, questioning the active management industry.

In the 1970s, John McQuown developed the first institutional index fund at Wells Fargo [1]. This marked the beginning of passive management and cap-weighted indexation. Index funds want to replicate the benchmark. On the contrary, active management intends to beat it. In recent years, the ETFs became the symbol of passive management, with their simplicity and low management costs, have had great success. However in the latest years, have caught on Style and Smart Beta ETFs, mixing features of both active and passive management, blurring their distinctions. In August 2021, Global ETF Assets hit 9 trillion [2]. Prominent asset managers such as BlackRock and Vanguard have significantly benefited from this trend. They increased assets under management by far in the latest years and mainly due to ETFs. Accordingly to Bloomberg, BlackRock is the largest ETF provider, with more than 3 trillion in assets under management, which represents one-third of the industry, as can be seen in fig. 4.1. Index funds' performances are measured with tracking error volatility (TEV). The lower it is, the better the index replicates the performances of the benchmark index.

$$TEV = \sqrt{\frac{1}{T-1}\sum_{t=1}^{T}\left(\tilde{r}_t - \tilde{r}_{B,t}\right)^2} \tag{4.1}$$

Where $\tilde{r}_t$ and $\tilde{r}_{B,t}$ are the returns of the fund and of the benchmark at time t, respectively. Exchange-traded funds can seek low TEV with a pure physical replication or by a synthetic one.

The weights on a Cap weighted index of *n-assets* are[3]:

$$w_{i,t} = \frac{N_{i,t}P_{i,t}}{\sum_{j=1}^{n} N_{j,t}P_{j,t}} \tag{4.2}$$

Where $N_{i,t}$ is the number of shares oustanding for the i-asset and $P_{i,t}$ is its price at time t. Cap weighted indexation is the most representative weighting of the market. Moreover, it is the only weighting scheme that is compatible with the efficient market hypothesis. Another feature of the CW portfolios is that they do not change if the market structure remains unchanged. The weight of an asset changes according to its return. They are the most convenient portfolios in terms of rebalancing and transaction costs.

[1] Roncalli, 2013, chap.1 p.18

[2] Wursthorn, 2021, Global ETF Assets Hit $9 Trillion

[3] Roncalli, 2013, chap.3 p. 154

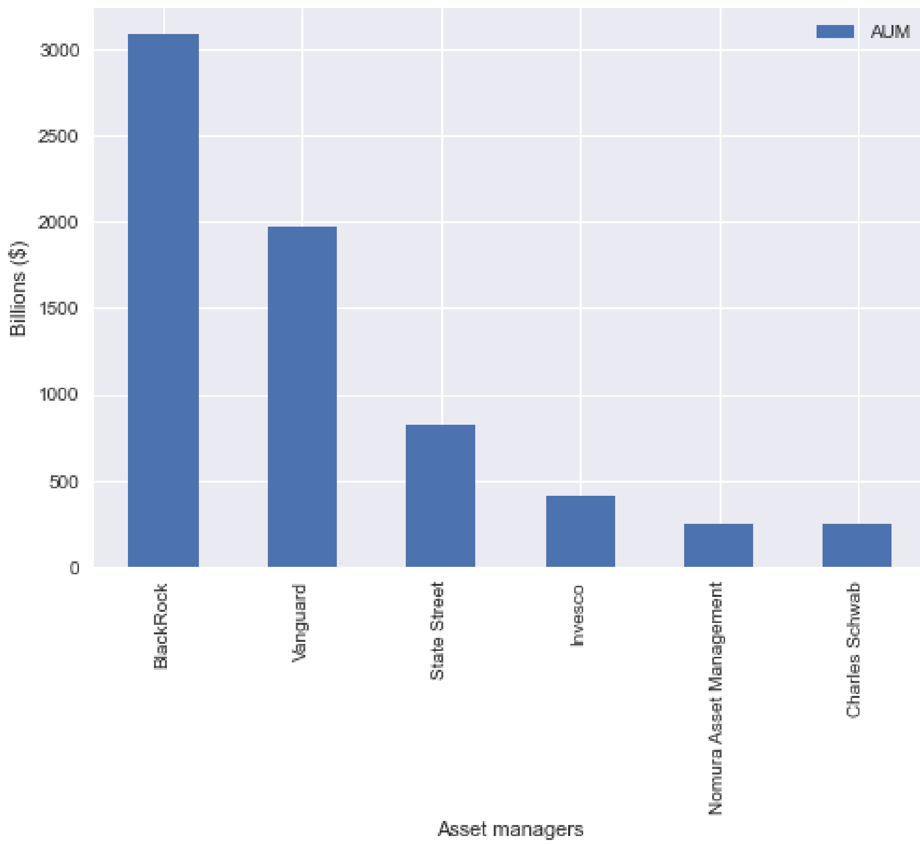

Figure 4.1: Asset managers ranked by ETFs under management

4.1.1 Inefficiency of Cap-weighted portfolios

Despite their popularity and simplicity, CW portfolios tend to be inefficient. They tend to be highly concentrated and poorly diversified, thus exposing the investor to specific risk. Accordingly to *slickcharts.com* data, the top 10 stocks of the S&P500 represent approximately 30% of the index. A team of professors from Cass business school,[4] wanted to test the inefficiency of such a weighting scheme. They selected 1000 companies over 43 years, and they programmed a computer to pick weights randomly. Thus, they simulated *the ability of stock picking of a "monkey"*. The process was iterated 10 million times. From these simulations, they found that many of the *monkey fund*

[4] Clare and Thomas, 2013

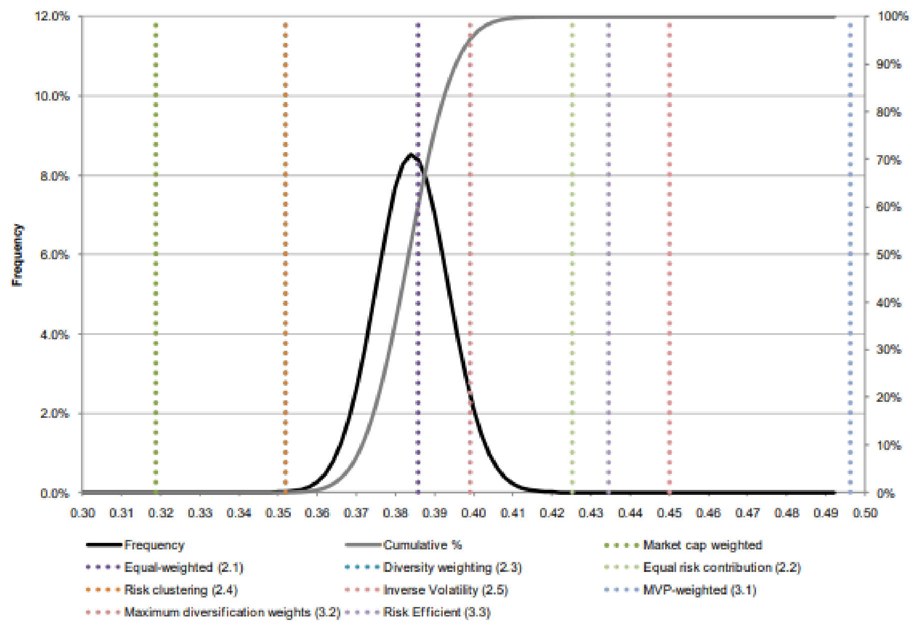

Figure 4.2: Distribution of monkeys' Sharpe ratios

managers would have generated superior performance of some alternative weighting schemes. The shocking result was that almost every monkey of the 10 million would have beat the performance of the cap-weighted index. Figure 4.2 below depicts the distribution of the monkey's Sharpe ratios. Even the values on the extreme left of the distribution are more significant than the Sharpe ratio of the CW indexation, which is 0.32[5].

Capitalization weighted portfolios tend to have a second drawback: they have an inefficient exposure to risk factors, regarding the Fama-French three-factor model, described in Chapter 2. It is empirically proven that small-cap stocks outperform large-cap stocks, and value stocks outperform growth stocks. Therefore, holding a cap-weighted means having a positive tilt towards value and large-cap stocks. We are allocating asset inefficiently regarding to SMB Fama-French risk factor. In addition, CW indexing is a trend-following strategy that incorporates momentum bias and leads to bubble exposure risk[6].

[5]the green dotted line to the extreme left
[6]Roncalli, 2013, chap 3 p.157

Therefore, the following sections aim to analyze alternative weighting schemes that may improve the disadvantages of the aforementioned portfolios.

4.2 Alternative weighting schemes

The goal of alternative indexing is to improve the risk-return profile of CW indices. This section reviews the most famous and used alternative weighting schemes, which tend to be more efficient than cap-weighting. As opposed to cap-weighted, these allocations will only be valid when the portfolio is established(t_0). Then, when assets prices change(t_1), the portfolio should be rebalanced. Active funds rebalance their portfolios just in a predetermined moment in time, for example, once every quarter or every year.

4.2.1 GMV portfolio

One of the commonest weightings for a risk-averse investor is the Global-Minimum Variance portfolio. The latter is designed to deliver the lowest possible portfolio variance, given a set of assets. Furthermore, it has an advantage over an optimized portfolio à la Markowitz: expected return estimates are not required. Instead, it requires just the covariance matrix as input. GMV portfolio corresponds to the following optimization problem:

$$\mathbf{w} = \min \frac{1}{2}\mathbf{w}^T \Sigma \mathbf{w} \quad u.c. \quad \mathbf{w^T 1} = 1 \tag{4.3}$$

where:

- $\mathbf{w}$ and $\mathbf{w}^T$ are nx1 and 1xn arrays with the initial weights.
- Σ is the covariance matrix.

GMV weights are then described as:

$$\mathbf{w} = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^T\Sigma^{-1}\mathbf{1}} \tag{4.4}$$

For this reason, even if it does not need expected return estimates, we are assuming them to be constant. While minimizing portfolio variance, this allocation gives up portfolio diversification since it will overweight low-volatility assets. This problem can be overcome in two ways:

1. imposing limits on the minimum number of assets.
2. assuming that volatility of each stock is equal and minimize portfolio variance given by $\text{Min}\,\sigma^2 \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \rho_{ij}$, thus creating the max-decorrelation portfolio.

However, GMV allocation presents good out-of-sample performance accordingly to Haugen and Baker (1991) and Clarke et al. (2006). Thereupon, it should be considered during the portfolio construction process when it is not possible to forecast returns with a high degree of accuracy.

4.2.2 Equally weighted portfolio

A naive approach to diversification suggests creating a portfolio well balanced in terms of dollar allocation. Therefore, the EW portfolio represents the most straightforward portfolio. For a portfolio of n assets, the weight of each asset will be $1/n$. Weights are defined based only on the number of assets, and stocks properties are not considered. This allocation is ideal if we assume that it is impossible to forecast risk and return. EW portfolio maximizes the Effective Number of Constituents:

$$ENC = \left(\sum_{i=1}^{N} w_i^2 \right)^{-1} \tag{4.5}$$

Where w_i represents the weight of each security and N is the number of assets. The ENC is a quantitative measure that describes how well diversified a portfolio is. In an equally weighted portfolio the $ENC = N$. EW portfolio coincides with the efficient portfolio if the returns and volatility are equal and correlation is uniform[7]. However, DeMiguel et al. (2009) shows the out-of-sample performances of the mean-variance optimization portfolios, such as the sample-based mean-variance, minimum-variance portfolio, value-weighted portfolio, and others with the naive 1/N portfolio as a benchmark. The article aims to understand the condition under which mean-variance optimized portfolios perform better than a naive EW strategy. The naive EW rule is used as a benchmark for its simplicity and, in addition, as Benartzi and Thaler (2001) states, despite sophisticated optimization models that have been developed in the latest years, investors continue to use such

[7] Roncalli, 2013, chap. 3 p.164

allocation. DeMiguel compares 14 different portfolio models across seven datasets of monthly returns. They showed that none of the models is consistently better than the EW strategy in terms of risk-return. These poor performances are mainly due to estimation errors. An optimized portfolio outperforms the benchmark if three conditions exist:

1. the estimation window is long.
2. the ex-ante (in-sample) Sharpe ratio is higher than that of the EW portfolio.
3. the number of assets is small.

A small number of assets implies fewer parameters to be estimated, therefore, a smaller estimation error. Moreover, fewer parameters lead to more negligible diversification in the EW scheme and, therefore, inefficiency. In particular, they discovered that the optimal length of the estimation window is 3000 months, in the case of monthly data (3000 data points) for 25 assets and more than 6000 data points for 50 assets. Regarding our analysis, our in-sample dataset, used to estimate parameters, covers five years with daily data. Considering 252 business days per year, we have approximately 1260 data points. Since we use 20 pairs, this will be the number of our assets. The number of data points in the in-sample period is quite large; however, it may give way to estimation error. For this reason, in our analysis, we will use the EW scheme as a benchmark, compared with a risk-based weighting scheme.

4.3 Risk budgeting

Risk budgeting does not allocate capital based on dollar allocations. Instead, it takes into account the risk of every asset and weights it accordingly. Weights are designed such that each asset contributes in a certain way to the portfolio's overall risk. Risk budgeting optimization is different from mean-variance optimization since, by construction, they do not consider expected returns. Even though the GMV portfolio does not require forecasts on returns, it implicitly assumes that they are constant. With the risk budgeting techniques, this assumption is not needed. Risk budgeting can implement active views of the fund manager to weigh a particular stock or asset class accordingly to the risk contribution wanted. We will focus on the equal risk contribution approach, where every asset has the same risk contribution to the portfolio's overall risk. This strategy in the industry took the name of Risk Parity.

4.3.1 Risk contribution

As we explained earlier, a well-balanced portfolio in terms of dollar allocation does not imply well balance of risk. To allocate our budget in terms of risk, we have to define the risk contribution of each asset. In our analysis, the risk is expressed as the portfolio variance:

$$\sigma_p^2 = \mathbf{w}^T \Sigma \mathbf{w} \tag{4.6}$$

Where $\mathbf{w}$ is an array with portfolio weights, and Σ is the covariance matrix. Then, the risk contribution of the i-assets is described as the portion of the portfolio's riskiness due to x.

$$RC_i = \frac{\mathsf{w}_i^2 \sigma_i^2 + \sum_{j \neq i}^{N} \mathsf{w_i w_j} \sigma_{ij}}{\sigma_p^2} \tag{4.7}$$

Where σ_{ij} represents the covariance between i and j asset. Let us make an illustrative example: we have a portfolio of two assets, *Hewlett-Packard*(HP) and *Advanced Micro Devices*(AMD)[8]. We construct the EW portfolio by allocating 50% of the budget each. This allocation will maximize ENC in formula 4.5. Thus, the overall risk of the portfolio is given by:

$$\sigma_p^2 = \mathsf{w}_{HP}^2 \sigma_{HP}^2 + \mathsf{w}_{AMD}^2 \sigma_{AMD}^2 + 2\mathsf{w}_{HP} \mathsf{w}_{AMD} \sigma_{HP,AMD} \tag{4.8}$$

HP risk contribution is described as:

$$RC_{HP} = \frac{\mathsf{w}_{HP}^2 \sigma_{HP}^2 + \mathsf{w}_{HP} \mathsf{w}_{AMD} \sigma_{HP,AMD}}{\sigma_p^2} \tag{4.9}$$

Where the last member of the equation 4.8 represents the covariance contribution divided by 2. Given the sample covariance matrix in table 4.1:

Risk contributions are shown in figure 4.3. As we can see from these values, approximately 80% of the risk is attributed to AMD, 20% to HP. So if we allocate our budget in terms of dollars, we will not set any constraints to the risk contribution of the assets.

[8] data provided by Bloomberg from January 2016 to January 2020

	HP	AMD
HP	0.000302	0.000227
AMD	0.000227	0.001674

Table 4.1: Sample covariance matrix.

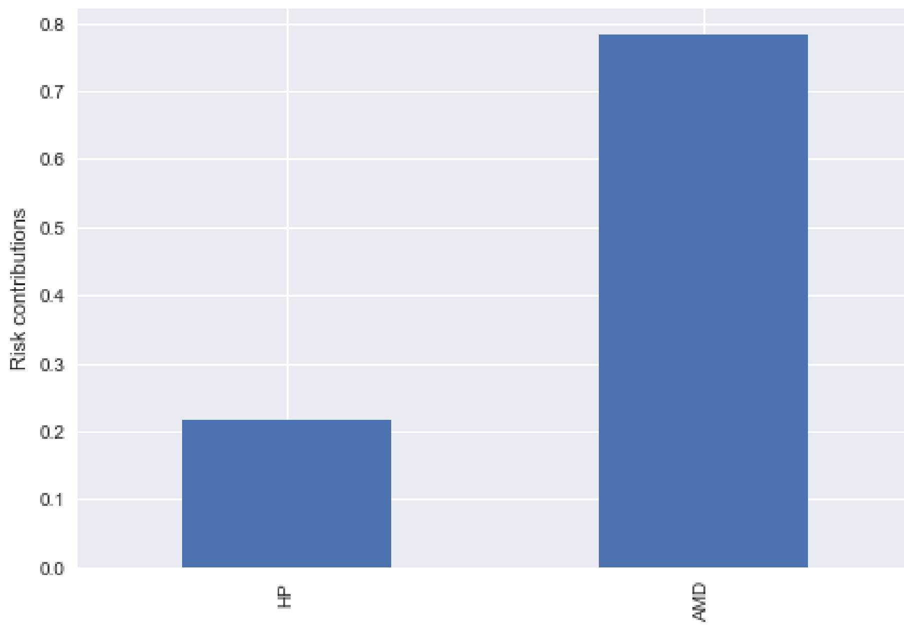

Figure 4.3: Risk contribution of HP and AMD EW portfolio.

4.3.2 Equal risk contribution

Risk parity portfolio is analogous to EW portfolio but in terms of risk allocation. The goal of the strategy is to set weights that make risk contributions equal among all assets. The target risk contribution of each security, described in equation 4.7 ,will be $1/n$. Since there is no analytical expression to find weights, we will use the *scipy* package with Least Squares Programming (SLSQP) to reach the target. The algorithm minimizes a function of several variables with any combination of bounds, equality, and inequality constraints. In our case, minimize the mean squared difference in risk contributions, given weights and a target risk contribution, that in the risk parity case, is $1/n$.

In the two assets case, risk parity weights are inversely proportional to the volatility of the component. This is confirmed from equation 4.10.

$$\begin{aligned} RC_1 &= \frac{\mathsf{w}_1^2\sigma_1^2 + \mathsf{w}_1\mathsf{w}_2\sigma_{1,2}}{\sigma_p^2} \\ RC_2 &= \frac{\mathsf{w}_2^2\sigma_2^2 + \mathsf{w}_2\mathsf{w}_1\sigma_{1,2}}{\sigma_p^2} \\ \mathsf{w}_1^2\sigma_1^2 &= \mathsf{w}_2^2\sigma_2^2 \\ \frac{\mathsf{w}_1}{\mathsf{w}_2} &= \frac{\sigma_2}{\sigma_1} \end{aligned} \tag{4.10}$$

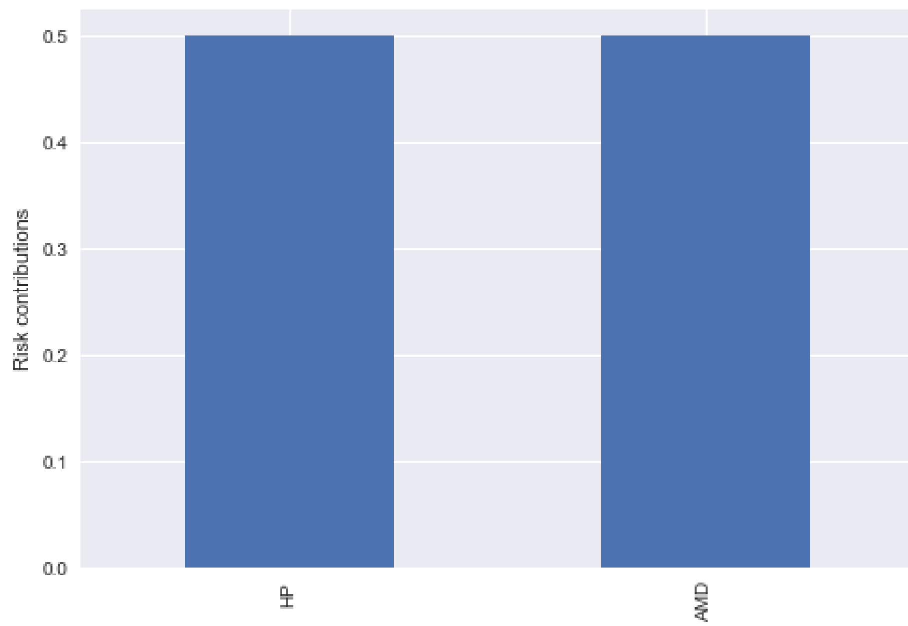

Figure 4.4: Risk contribution with risk parity weights

Considering the previous section's example, we expect to weigh more HP and less AMD to achieve equal risk contribution. Indeed, applying the optimizer will return 70% for HP and 30% for AMD. In figure 4.4 we can see that risk contributions are now equal.

Risky parity portfolios tend to overweight less risky components but less strongly than GMV portfolios. As a result, they are more risk efficient than the abovecited benchmarks. In the following sections, we will describe the implementation of Risk Parity in the trading strategy.

4.4 Implementation

Typically, risk parity is used to allocate risk among asset classes. So, for example, if we want to create a portfolio with equity, bonds, and commodities, we can find the proper budget to invest in each asset class.

In our analysis, we propose its implementation in a trading strategy. First, we will define *synthetic* assets, pairs will be treated as a single asset, and then we will construct the covariance matrix.

The pairs trading process involves trading two stocks simultaneously, opening a long and a short position, and closing them together. We will consider the total return of the positions, for each pair, as the return of each synthetic asset. In other words, we are assuming that the pair is an asset that generates return[9] while positions are opened.

In our review, we consider 20 pairs, therefore 20 time series of returns. We will use the equal risk contribution principle to determine the initial weighting of each pair in the total trading portfolio. In particular, we will compute the return of each pair regarding the in-sample period and then construct the covariance matrix. Since it is in-sample backtesting, returns will suffer from a bias due to the usage of information from the future. However, this bias will not affect our result excessively since we aim to weigh pairs based on the variance of returns, not their performance. Once we have the covariance matrix, we will compute the risk contribution of each asset, and then we will apply the risk parity weighting. We will assume an equally weighted portfolio among pairs. Figure 4.5 depicts the risk contribution of every pair.

It is clear that, in our case, a well-balanced allocation in terms of dollars does not imply a balanced risk among assets. Just five pairs contribute to approximately 50% of the portfolio's overall risk. In particular, the *synthetic* asset composed by *Cintas* and *Expeditors International of Washington Inc.* contribute the most, and the pair formed by *American Electric Power* and *Eversource Energy* has a risk contribution of around 0.3%. Returns well confirm this, as can be seen in figure 4.6. If we assume these two assets as two different funds managed by two fund managers, they would have two distinctly different risk profiles. Inclined to risk, the former would experience severe drawdowns, the latter risk-averse, with a maximum drawdown of 3%. In our optimization, we expect the first pair to reduce his weighting and the second to increase it.

[9] further information about the computation of returns will be given in the next chapter.

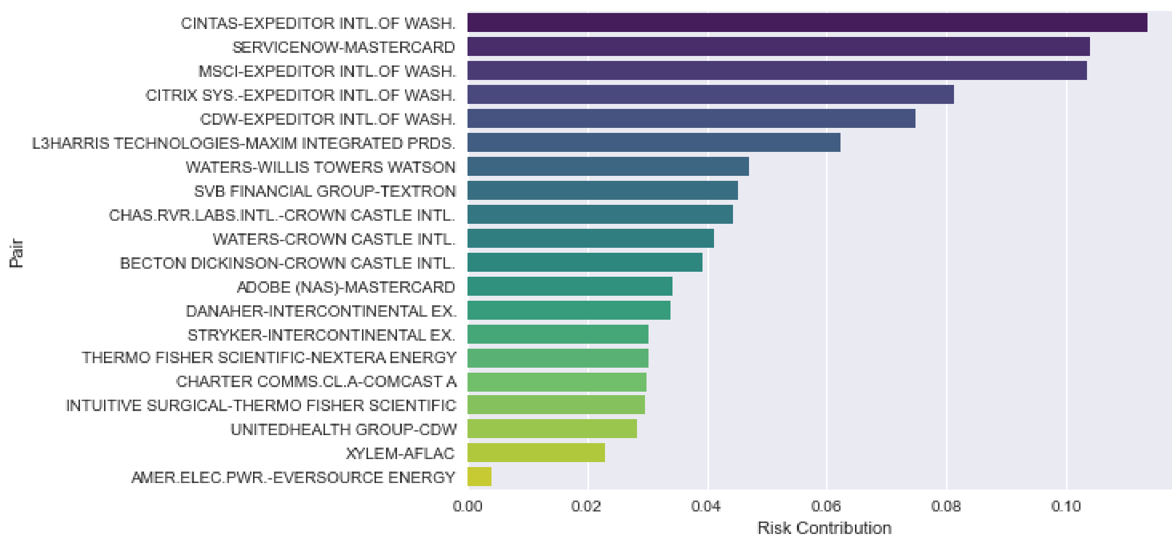

Figure 4.5: Risk contribution of each pair in the EW portfolio

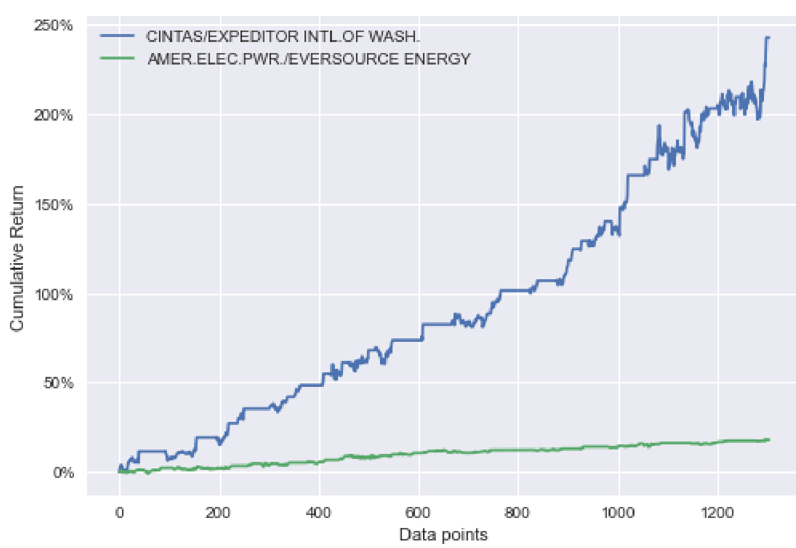

Figure 4.6: Cumulative return of extreme contribution assets

Pair	ERC Weight
CDW/EXPEDITOR INTL.OF WASH.	0.034969
XYLEM/AFLAC	0.065689
CHAS.RVR.LABS.INTL./CROWN CASTLE INTL.	0.041461
CINTAS/EXPEDITOR INTL.OF WASH.	0.024317
SERVICENOW/MASTERCARD	0.027763
ADOBE (NAS)/MASTERCARD	0.051315
INTUITIVE SURGICAL/THERMO FISHER SCIENTIFIC	0.050398
SVB FINANCIAL GROUP/TEXTRON	0.041907
BECTON DICKINSON/CROWN CASTLE INTL.	0.048026
UNITEDHEALTH GROUP/CDW	0.054849
THERMO FISHER SCIENTIFIC/NEXTERA ENERGY	0.052649
STRYKER/INTERCONTINENTAL EX.	0.050688
WATERS/WILLIS TOWERS WATSON	0.042410
CHARTER COMMS.CL.A/COMCAST A	0.055217
AMER.ELEC.PWR./EVERSOURCE ENERGY	0.165805
CITRIX SYS./EXPEDITOR INTL.OF WASH.	0.032699
MSCI/EXPEDITOR INTL.OF WASH.	0.027163
WATERS/CROWN CASTLE INTL.	0.046226
L3HARRIS TECHNOLOGIES/MAXIM INTEGRATED PRDS.	0.036596
DANAHER/INTERCONTINENTAL EX.	0.049850

Table 4.2: Risk parity weights of each pair based on in-sample parmeters

Indeed, their weights will become 2.42% and 16.5%, respectively, rather than the 5% in the EW portfolio. We will set ERC weights, displayed in table 4.2, at the beginning of our out-of-sample testing period, 2019Q1. Since the validation period is two years and a half, to make the most of this asset allocation, we will not rebalance the weights. With this weighting scheme, we should have a more balanced portfolio in terms of risk and, therefore, performs better in the backtesting. We will test this allocation, comparing it with the EW portfolio, in the next chapter.

Chapter 5

Performance Analysis

This chapter will analyze the performance of the trading strategy regarding the out-of-sample period (2019Q1-2021Q2). We will backtest the trading strategy with two different weightings. In particular, we will compare the performances of an EW portfolio among pairs, and a Risk parity one. In the first case, each *synthetic* asset's initial budget will be equal, in the second one will be based on the risk contribution.

5.1 Return Computation

Since the focus of the performance is on returns, we first need to explain how they are calculated. To better understand the whole process, let us summarize our trading strategy.

1. Pairs are selected.
2. We construct the spread series.
3. When spread series deviate from equilibrium, thresholds are triggered, we open positions on both assets composing the pair simultaneously. In particular:
   - we open a position of 1 on *Y-asset*.
   - we open a position of β on the *X-asset*.
4. When the spread reverts to equilibrium, positions are closed.

This process is done for 20 pairs, therefore, 40 assets. Our goal is to construct the P&L, the total profit of the strategy for every t. Initially, we want to create an equally weighted portfolio among pairs. To design such a trading portfolio, we need to trade CFDs[1]. We need the stocks to be partitionable to assign the same budget to each pair, regardless of the stock prices. In truth, this hypothesis does not depart from reality, since evaluating the implementation of this algorithm, many brokers who support API calls use these derivatives. Moreover, brokers encourage the trading of these derivatives, allowing retail investors to enter the market with little capital.

5.1.1 Estimation process

The process of return estimation will be as follows:

1. computes individual return of each asset day by day.
2. weigh such return for the weight of the asset in the pair, given the quantity constraints.
3. weigh the return referred to in point 2 for the weight attributed to that particular pair, in our case: EW and Risk Parity weights.

Asset return

To compute the return of every position, we will use the methodology of Gatev et al. (2006). We expect the pairs to open at different times during the quarter trading periods since we selected pairs with the lowest half-life. As we mentioned in section 3.3, positions opened will be closed on the last trading day of the quarter. Profits and losses are computed over long and short positions of 1\$; therefore, they will be interpreted as excess returns. The excess return on a pair during a trading interval is computed as the reinvested payoffs during the trading interval. Marking-to-market is daily. Daily returns of long-short positions are calculated as follows.

$$\begin{aligned} long : r_{i,t} &= sr_{i,t}w_{i,t} \\ short : r_{i,t} &= -sr_{i,t}w_{i,t} \end{aligned} \tag{5.1}$$

[1]CFDs are derivatives that pays the differences in the settlement price between the open and closing trade, they allow investors to bet on price movements

where $w_{i,t}$, the wealth level of the $i - asset$ at time t is defined as:

$$w_{i,t} = w_{i,t-1}(1 + r_{i,t-1}) = \prod_{t=1}^{t-1}(1 + r_{i,t}) \tag{5.2}$$

where $r_{i,t}$ represent the return of $i - asset$ at time t taking into account reinvested pay-off and $sr_{i,t}$ represents the simple return of $i - asset$, obtained with the percentage change of prices between t and $t - 1$.

Portfolio returns

Since we have quantity constraints, unlike Gatev et al. (2006) analysis, we cannot set the weight of Y and X assets in the pair to 50% each. Therefore, the weighting of the components of a pair at time t depends on prices. Furthermore, we need to weigh the return by the number of pairs. In particular, given a pair of Y and X, weights within the pair are described in equation 5.3 and 5.4, respectively:

$$q_{y,t} = \frac{p_{y,t}}{p_{y,t} + \beta p_{x,t}} \frac{1}{n} \tag{5.3}$$

$$q_{x,t} = \frac{\beta p_{x,t}}{p_{y,t} + \beta p_{x,t}} \frac{1}{n} = (1 - q_y)\frac{1}{n} \tag{5.4}$$

where $p_{y,t}$ and $p_{x,t}$ are the prices of Y and X assets at time t, β represents the cointegration coefficient and n is the number of pairs. These weights remain constant while the position is opened. However, when the position is closed and reopened, the weights are updated to capture price changes. The individual return at every t is weighted with q_t , to get to *intrapair* returns. The return over the portfolio of asset j at time t will be:

$$rp_{j,t} = \frac{q_{j,t} sr_{j,t} w_{j,t}}{\sum_{j \in P} q_{j,t} w_{j,t}} \tag{5.5}$$

And the total return of the portfolio at time t is described as:

$$r_{P,t} = \sum_{j \in P} rp_{j,t} \tag{5.6}$$

Since we want to test an alternative weighting scheme, we will change the $\frac{1}{n}$ in equation 5.3 and 5.4 with the weight attributed to each pair.

5.2 Strategy returns

This section will review the strategy performance, provide the returns, risk-adjusted measures, and analyze the strategy's overall riskiness. The strategy performed well in both EW and ERC portfolios. In particular, we generated a net return of 156% over two and a half years in the EW portfolio. The risk parity portfolio generated a net return of 143%. Figure 5.1 depicts the cumulative return of the strategy. Such returns annualized become 43% for

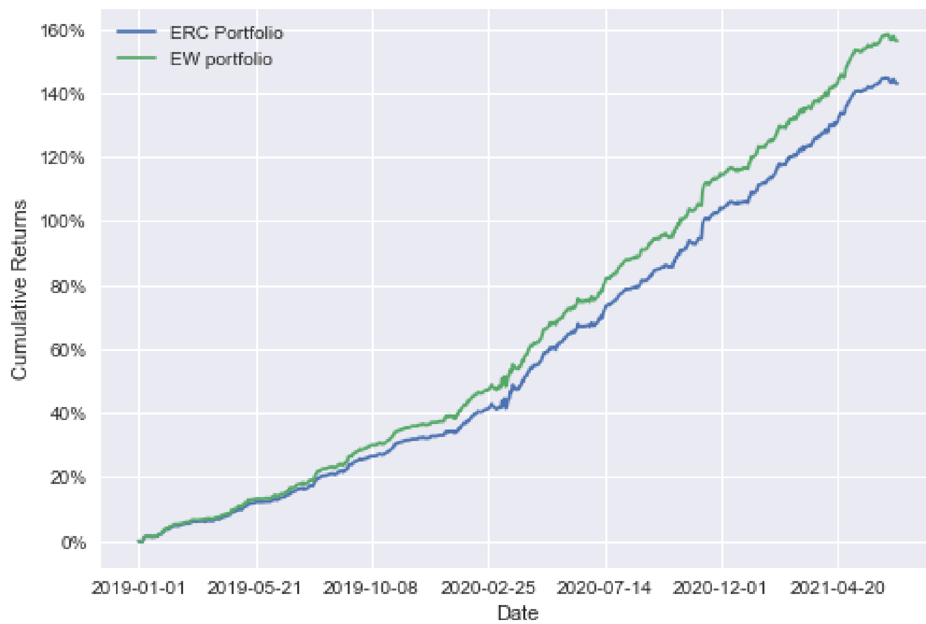

Figure 5.1: Cumulative returns of EW and ERC portfolios

the EW and 40% for the ERC, as can be seen in table 5.1.

	Annualized Return	Annualized Volatility	Maximum Drawdown
EW	0.438599	0.050172	-0.020758
ERC	0.409259	0.047062	-0.021270

Table 5.1: Summary statistics of portfolio returns

5.2.1 Pairs characteristics

In this section, we will analyze how pairs contributed to the portfolio return. The best performing pair has been *Servicenow - Mastercard*. It generated a return of 343% over his initial budget. On the other hand, the worst pair in terms of returns has been *American Electric Power -Eversource Energy*, with a 33%. Hence, the former has outperformed the latter by ten times. This confirms the robustness of the analysis; pairs are keeping their behavior out-of-sample. In particular *Servicenow-Mastercard* was the second-best performing pair, and *American Electric Power-Eversource* was the worst-performing one, regarding the in-sample analysis.

The difference in performance may be due to the industry difference. We notice that the component of these two pairs belongs to the same sector. This is because of the co-movement we were looking for in pair selection. Furthermore, their cointegration coefficient is approximately 1. In other words, a difference between their prices generates a stationary process. The former is a "tech" pair and the latter an "energy" one. It is known that the tech sector can be pretty risky but, on the other hand, can secure higher returns. Meanwhile, the energy industry tends to be less volatile.

The number of trades confirms this, indeed as can be seen in figure 5.2, "tech" and "energy" pairs have been traded 64 and 58 times, respectively. Since the number of trade is approximately the same, each transaction on the "tech" secured a much higher profit than the latter. In our case, evidence suggests that higher volatility of the traded assets brings higher returns. Broader movement in prices may give way to higher returns.

Regarding the EW portfolio, pairs that contribute the most to the overall return are the pairs that performed better. However, this is not true in the ERC portfolio since weights are based on in-sample risk contribution. Hence, even though a pair generates an abnormal return, it can be weighted less than $1/n$[2] due to its risk contribution. In particular, in our case, the most performing pairs are underweighted, apart from one specific pair: *UnitedHealth Group - CDW Corporation*. This pair is optimal, has a low risk contribution, and generates an out-of-sample return of 292%. Applying the ERC its weight on the portfolio rises; therefore, its contribution to the portfolio return in the Risk parity portfolio is higher than in the EW. However, this is not sufficient to balance the effects of the weight reduction in high performing assets, as

[2] most of the times

can be seen in table5.2.

We are increasing the weight of just one out of five, pairs. In addition, the risk contribution of *American Electric Power - Eversource Energy* is 0,03%. We can consider it an outlier. The ERC scheme overweight such asset to 16%, leaving no room for optimal cases as *UnitedHealth Group - CDW Corporation*.

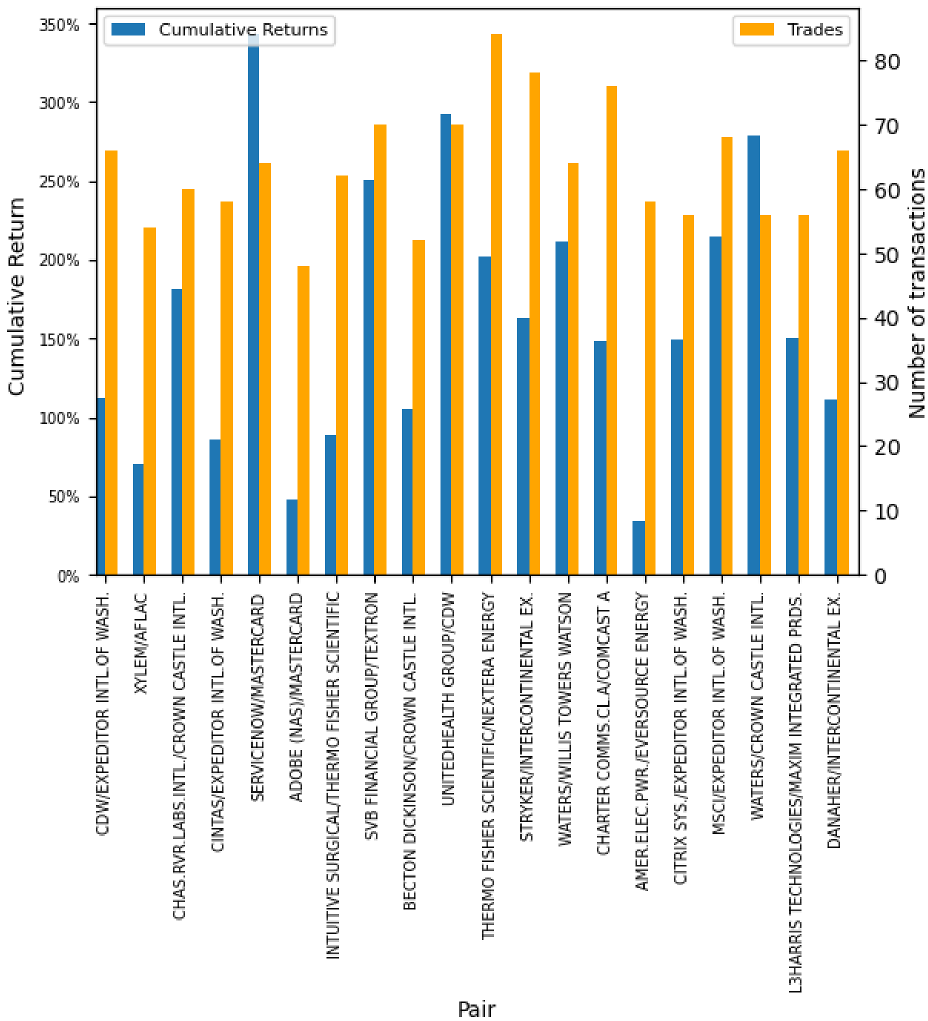

Figure 5.2: Cumulative returns compared with number of trades

Pair	Risk Contribution	W_{EW}	W_{ERC}
SERVICENOW-MASTERCARD	0.103878	0.05	0.027763
UNITEDHEALTH GROUP-CDW	0.028330	0.05	0.054849
WATERS-CROWN CASTLE INTL.	0.041260	0.05	0.046226
SVB FINANCIAL GROUP-TEXTRON	0.045154	0.05	0.041907
MSCI-EXPEDITOR INTL.OF WASH.	0.103491	0.05	0.027163

Table 5.2: Change in weights of the 5 most performing asset between EW and ERC scheme

5.2.2 Alternative risk measures

The most common measure of risk is volatility. In our portfolios, annualized volatilities are 5% for the EW and 4.7% for the ERC. Given such high returns, we can assume that regarding to this risk measure, our strategy is solid. However, there are other risk measures to consider.

Semi-deviation

For a risk-averse investor, the simple variability of returns may not be enough. In particular, volatility takes into account positive and negative fluctuations. Since positive shifts of returns are desired, semi-deviation considers just unfavorable fluctuations of returns. It is described in the equation 5.7:

$$Semi - deviation = \sqrt{\frac{1}{n}\sum_{r_t<\overline{r}}^{n}(\overline{r} - r_t)^2} \tag{5.7}$$

Where $\overline{r}$ represent the mean of the return series, r_t is the return at time t and n is the total number of observation below the mean.

In our case, we have an annual semi-deviation of 4.2% for the EW and 3.8% for the ERC. We were expecting this result since the risk parity is designed to be more balanced in terms of risk; it should reduce the overall riskiness, and therefore, the semi-deviation. If the investor's goal is to preserve his capital, he should pick the ERC portfolio. However, this may be contradicted by drawdowns.

Maximum Drawdown

An essential indicator of the downside risk is maximum drawdown. It is described as the maximum observed loss from a peak to a trough of a portfolio. Drawdown at time t is defined as:

$$DD_t = \frac{TroughValue_t - PeakValue}{PeakValue} \tag{5.8}$$

Risky portfolios tend to have higher drawdowns than "non-risky" portfolios. In our case, EW represents the naive diversification in terms of risk, and therefore, the risky portfolio. ERC, on the other hand, represents the "conservative" portfolio.

Figure 5.3 depicts drawdowns. We can see that they are pretty similar. However, we have an unexpected outcome. The maximum drawdown is higher in the ERC portfolio, -2.12% compared to -2.07%. Even though they differ only by five basis points, this result proved that the "conservative" weighting scheme suffered market turbulence. Indeed maximum drawdowns in both portfolios manifested the same day, March 17, 2020. That day financial markets were in the middle of the covid crisis. In particular, we can find an analogy between the performance of S&P500 and our strategy. The maximum drawdown of S&P500 occurred on March 23, 2020, with a -34%.

Furthermore, if we look at the VIX index, also known as the *fear index*, which represents the stock market's expectation of volatility based on S&P 500 index options, we can see that on March 16, it surged around 43%, closing to a record high of 82.69, surpassing for the first time the peak level of November 21, 2008.

Our strategy aims to market neutrality. However, evidence suggests that there is a minimum relationship between S&P500 and our portfolio returns. Nonetheless, our strategy suffered losses of 2%, compared to 34%. If we consider two investors:

- A: long-only on S&P 500
- B: investing in our strategy

In March 2020, Investor A would have born a loss 17 times superior to investor B. Thus, the strategy produces excellent drawdown results. On the other hand, the ERC weighting scheme fails, suffering a higher drawdown than the EW portfolio.

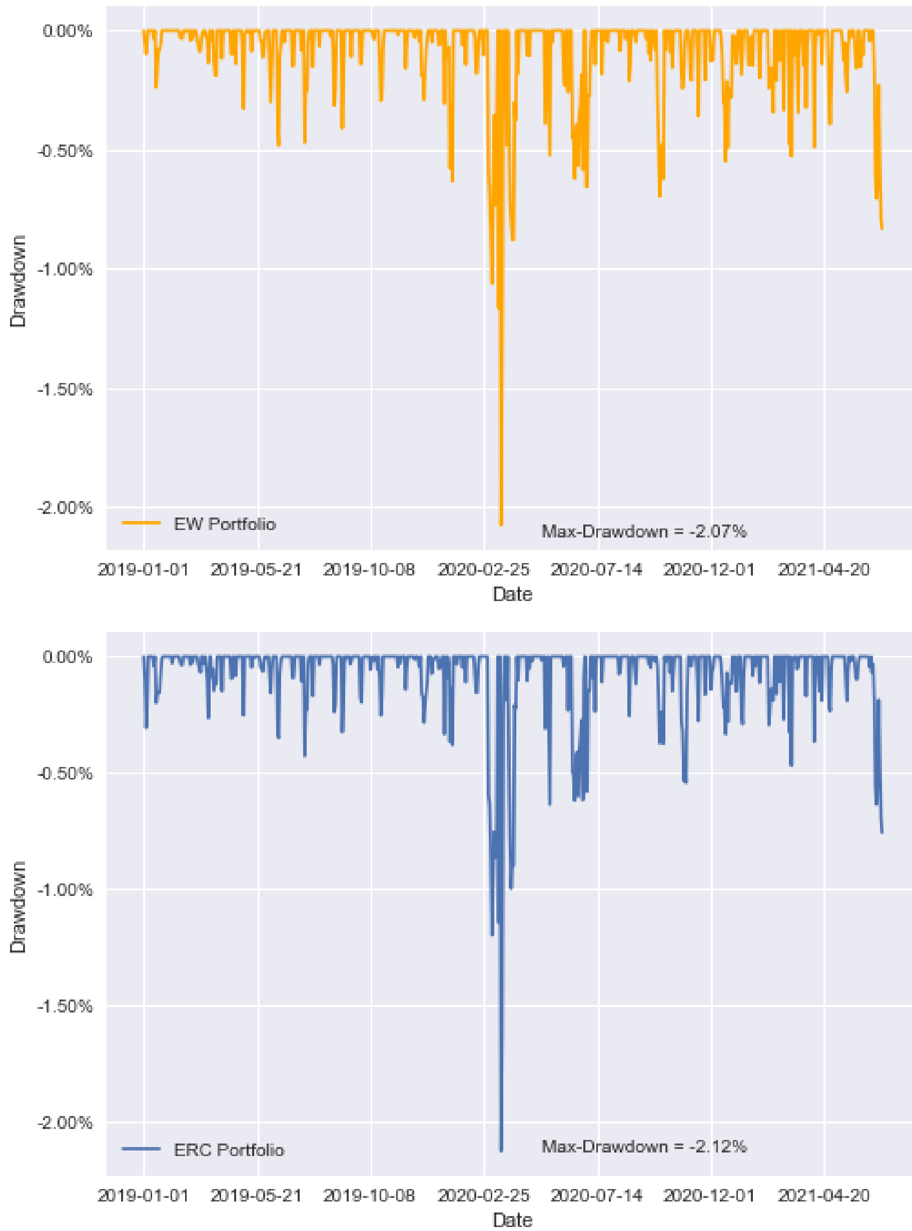

Figure 5.3: EW and ERC portfolios drawdowns

Strategy's Beta

We want to analyze the relationship, between S&P500 and the returns of our strategy, that came out of drawdown measures. In section 2.2 we described long-short strategies as market neutral. In particular, we want to check if the beta of our strategy is negligible. Therefore, we will perform a regression between the market excess returns and the excess return of our portfolios. Results of the regression are displayed in table 5.3 and 5.4.

We notice that the betas of both portfolios are close to 0. In particular, given p-values, we see that both betas are not significant. Results show market neutrality. In EW and ERC portfolios, beta cannot explain returns, proving they are uncorrelated with the market. At the same time, Jensen's alpha is positive and significant. They are 0.031 and 0.029, respectively. Such high alphas state that the unexplained returns are approximately 3,1% and 2.9% monthly. Which annualized represents the total return of the strategy. Outcomes have shown that the strategy is indeed market neutral.

EW	Coefficient	Std. Err.	P-val.
Beta	0.0128	0.045	0.780
Alpha	0.0315	0.003	0.000

Table 5.3: EW portfolio regression

ERC	Coefficient	Std. Err.	P-val.
Beta	0.0304	0.044	0.494
Alpha	0.0293	0.003	0.000

Table 5.4: ERC portfolio regression

5.2.3 Risk-Adjusted measure

It is well known that exists a trade-off between risk and return. If we evaluate funds' performance by just looking at returns, we make a mistake. We need to correct returns with a risk measure. The most famous ratio in such a sense is the Sharpe ratio, defined in equation 5.9:

$$SR = \frac{R_p - R_f}{\sigma_p} \tag{5.9}$$

Where R_p is the return of the portfolio, R_f is the risk-free and σ_p represents the standard deviation of the portfolio's excess return.It represents the reward per unit of risk.

In our analysis, the EW portfolio performed better than the ERC in terms of returns. Theoretically, we expected such a result. Therefore, the $1/n$ portfolio can generate higher returns. ERC principle aims to optimize risk allocation, not achieving abnormal returns. Clearly, by constructing the risk parity portfolio, we are giving up returns to achieve better risk-adjusted-performance. However, we can see from figure 5.1 that the returns are pretty similar. Indeed annualized volatilities, as can be seen in table 5.1, are 5% and 4.7%, respectively. Thus, we give up a 3% percent annual return to gain a -0.3% portfolio volatility. Given such low volatilities for both portfolios, it's not worth renouncing on a 13% return to save 1% of riskiness. This, as can be seen in table 5.5, is confirmed by the Sharpe ratio, which is higher in the EW portfolio rather than the ERC.

	EW portfolio	ERC portfolio	S&P500 index
Sharpe Ratios	8.720457	8.673775	0.855131.

Table 5.5: Sharpe ratios of different portfolios

The risk parity scheme should have outperformed the naive allocation in terms of Sharpe ratio but reduced return more than reducing risk. In other words, results have stated that the ERC portfolio is inefficient.

On the other hand, Sharpe ratios are huge compared to market Sharpe ratios. In particular, the Sharpe ratio of the strategy is approximately ten times bigger than the one in the market. We cannot compare a passive buy-hold strategy with an active one. Even though we compare it with the active management industry Sharpe ratios, a value of 8 is really high. These superior performances will be justified in the last section.

5.2.4 EW vs. ERC

The strategy has been backtested with two different weighting schemes. Naive diversification, where every pair weights $1/n$, and Risk parity, where pairs are weighted to achieve equal risk contribution. The second scheme relies on in-sample returns to better allocate assets in terms of risk contribution.

The ERC portfolio, in our strategy, turned out to be inefficient. It produced worse results than the EW portfolio in terms of returns, drawdown, and risk-adjusted return. Figure 5.4 depicts the ex-post risk contribution of the portfolio in both weighting schemes. We can see that in the risk parity, there are fewer outliers. However, as we said, this allocation, even if more balanced in terms of risk, did not improve performances.

In particular, in the EW, the "asset" *Servicenow- Mastercard* has a risk contribution more significant than the others. However, this is probably due to the upside rather than downside risk. The risk parity inefficiency can be caused by the outlier *American Electric Power -Eversource Energy* that, overweighted, may not leave space to other assets.

Our analysis infers that ERC is not optimal to implement in a statistical arbitrage trading strategy. Hence, given the low risk of the strategy itself, if an investor has to allocate his budget among pairs, results suggest that creating an equally weighted portfolio is a much better solution.

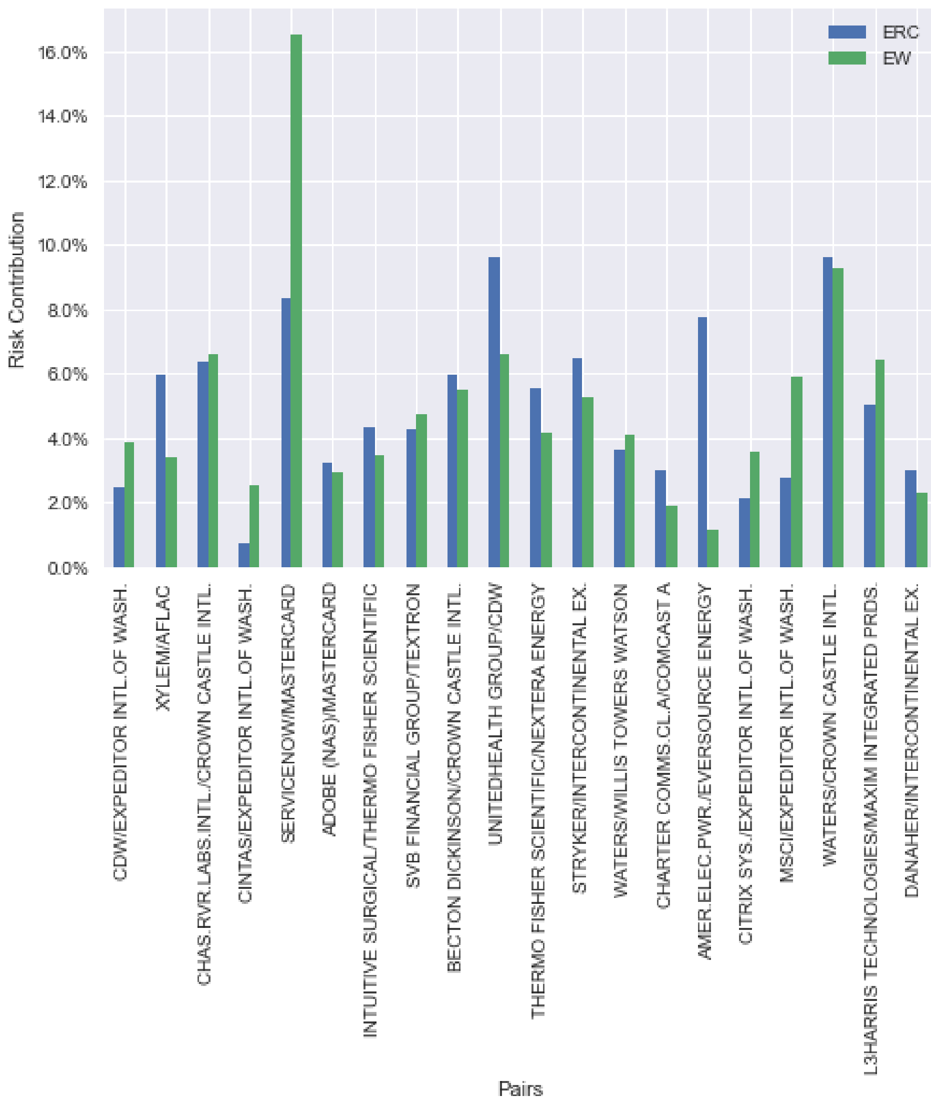

Figure 5.4: Ex-post risk contribution

5.2.5 All that glitters is not gold

This section will analyze the issues and the hidden costs of the trading strategy. First of all, our analysis does not consider transaction costs. Since we invest in an active strategy, there will be a lot of transactions. In particular, we executed a total of 1266 trades on the pairs over two years and a half. Gatev et al. (2006), states that transaction costs are around 50 basis points. If we compute a naive estimation of transaction costs, we can arrive at a total cost of 30%. Leading annualized return to 32% for the EW and 29% for the ERC portfolio. Even though this metrics highly influences returns, the Sharpe ratio, would be around 6. In addition, there are hidden costs that we are not taking into account. Most of them are related to short-selling. In particular, lending fees, stop-loss margins, and margin calls.

Lending fees can be daily; for this reason, positions opened for too long may be less profitable. Over 20 pairs, our positions remained open for approximately 400 days per pair. Since it is a long short-strategy, this means a total of 400 days short in each pair.

Margin calls occur when the value of an investor's margin account falls below the broker's required amount. As a result, the investor is obliged to post more collateral or close the position, reducing the strategy's profitability again.

The validation period covers two years and a half. So even though there is a financial crisis in the middle of it that tests the robustness of our market neutral strategy over different market conditions, the period may not be long enough to assert the absolute profitability of the strategy.

However, the performance of the strategy was excellent. For this reason, taking into account these costs will inevitably reduce the returns but will not invalidate the results.

Chapter 6

Conclusions

In this thesis, we introduced the concepts of Pairs trading and long-short equity strategies. We reviewed stationarity, cointegration, and market neutrality. In addition, we presented the commonest weighting schemes, emphasizing their benefits and drawbacks, comparing them with risk budgeting techniques.

As thesis core, we developed an algorithm that backtests a statistical arbitrage trading strategy. In particular, it tests for cointegration relationships among a broad set of assets and selects the most tradable pairs. Subsequently, it creates trading signals to enter or exit a trade. Finally, it computes the daily P&L and summary statistics of the portfolio.

Furthermore, we proposed a new weighting scheme for a statistical arbitrage portfolio that relies on the equal risk contribution principle. Such allocation aims to enhance the risk-adjusted performance of the portfolio, reducing its overall risk.

The algorithm tested the strategy in the US market regarding the period 2014-2021, with two years and a half of out-sample backtesting. The trading strategy generated an abnormal return of 40% annualized, with extremely low drawdowns and high risk-adjusted measures. These outcomes have to be seen through the lens proposed in section 5.2.5. In addition, results showed that our returns are uncorrelated with the market. Therefore, the strategy has been successful in achieving market neutrality.

However, the risk parity weighting scheme did not achieve its purpose. The equally weighted portfolio outperformed the equal risk contribution portfolio in terms of returns, maximum drawdown, and Sharpe ratio. We conclude that given the low risk of the strategy, a more balanced allocation in terms of risk generates inefficiencies. Results suggest that we are exposed to upside risk more than downside risk. For this reason, balancing assets inversely proportional to their risk contribution ends up reducing returns rather than limiting risk.

In conclusion, within our analysis, the strategy performed well. Given such promising results, this strategy will be improved and tested in the real market.

Bibliography

S. Benartzi and R. H. Thaler. Naive diversification strategies in defined contribution saving plans. *American economic review*, 91(1):79–98, 2001.

Z. Bodie et al. *Investments*. Tata McGraw-Hill Education, 2009.

M. Clare and Thomas. An evaluation of alternative equity indices part 1: Heuristic and optimised weighting schemes. 2013.

R. G. Clarke, H. De Silva, and S. Thorley. Minimum-variance portfolios in the us equity market. *The journal of portfolio management*, 33(1):10–24, 2006.

V. DeMiguel, L. Garlappi, and R. Uppal. Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, 22(5):1915–1953, 2009.

R. F. Engle and C. W. Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.

E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst. Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, 19(3):797–827, 2006.

R. A. Haugen and N. L. Baker. The efficient market inefficiency of capitalization–weighted stock portfolios. *The journal of portfolio management*, 17(3):35–40, 1991.

F.-S. Lhabitant. *Handbook of hedge funds*, volume 332. John Wiley & Sons, 2007.

T. Roncalli. *Introduction to risk parity and budgeting*. CRC Press, 2013.

G. Vidyamurthy. *Pairs Trading: quantitative methods and analysis*, volume 217. John Wiley & Sons, 2004.

M. Wursthorn. Global etf assets hit $9 trillion, 2021. URL https://www.wsj.com/articles/global-etf-assets-hit-9-trillion-11628769548.

Appendices

.1 Code

```python
def find_cointegrated_pairs(data,sortby = "half_life"):
    n = data.shape[1]
    score_matrix = np.zeros((n,n))
    pvalue_matrix = np.ones((n,n))
    keys = data.keys()
    pairs = [] # just an empty list that has to be define before the for cicle
    score_pairs=[]
    pval_pairs = []
    for i in range(n):
        for j in range(i+1,n):
            Series1 = data[keys[i]]
            Series2 = data[keys[j]]   # with this i isolate the two series from the data frame
            result = coint(Series1, Series2)
            score = result[0]
            pvalue = result[1]
            score_matrix[i,j] = score  # with this i subsistute in the all 0 matrix my correct score for the pair
            pvalue_matrix[i,j] = pvalue  # with this i subsistute in the all 1 matrix the correct p-value for the pair
            if pvalue < 0.05:
                pairs.append((keys[i],keys[j]))
                score_pairs.append(score)
                pval_pairs.append(pvalue)

    hfs,pva = halflife_pairs(data,pairs) # i compute the halflife

    s = np.array(score_pairs)
    s_n = (s-np.mean(s))/np.std(s)
    h = np.array(hfs)
    h_n = (h-np.mean(h))/np.std(h)
    skore = s_n + h_n

    ranked_pairs = pd.DataFrame(data={"pair":pairs,"score":skore,"half_life":hfs,"score_test":score_pairs,"significance":pva}).sort_values(by=sortby)
    drop_list = ranked_pairs.loc[ranked_pairs.significance>0.05].index.to_list()
```

```python
    ranked_pairs.drop(index=drop_list,inplace=True)
    ranked_pairs.reset_index(drop=True,inplace=True)
    return score_matrix, pvalue_matrix, pairs, ranked_pairs
```

```python
def Coint_coeff(data,ranked_pairs, n_pairs):
    '''
    run two regression and store the params with higher beta
    '''
    betas = []
    y_s = []
    x_s =[]
    pva = []
    pairs = ranked_pairs.loc[:n_pairs,"pair"].tolist()
    for i in range(len(pairs)):
        #first regression
        Y = data.loc[:,pairs[i][0]]
        X = data.loc[:,pairs[i][1]]
        X = sm.add_constant(X)
        model_1= sm.OLS(Y.astype(float),X.astype(float))
        results_1 = model_1.fit()
        beta_1 = results_1.params[1]
        p_val_b1 = results_1.pvalues[1]
        #second regression
        Y = data.loc[:,pairs[i][0]]
        X = data.loc[:,pairs[i][1]]
        Y = sm.add_constant(Y)
        model_2= sm.OLS(X.astype(float),Y.astype(float))
        results_2 = model_2.fit()
        beta_2 = results_2.params[1]
        p_val_b2 = results_2.pvalues[1]
        #if significant
        if beta_1>beta_2:
            betas.append(beta_1)
            y_s.append(pairs[i][0])
            x_s.append(pairs[i][1])
            pva.append(p_val_b1)
        if beta_1<beta_2:
            betas.append(beta_2)
            y_s.append(pairs[i][1])
            x_s.append(pairs[i][0])
            pva.append(p_val_b2)

    return betas, y_s, x_s, pva
```

```python
def spread_mat(data, y_s, x_s, betas):
    '''
    return the matrix of the spread series
    '''

    spread_matrix = np.zeros((len(data),len(y_s)))

    for i in range(len(y_s)):
        x_b = data.loc[:,x_s[i]]*betas[i]

        y = data.loc[:,y_s[i]]
        spread_matrix[:,i] = y-x_b


    names =[]
    for i in range(0,len(x_s)):
        y_x = y_s[i]+"/"+x_s[i]
        names.append(y_x)

    spread_df = pd.DataFrame(data=spread_matrix)
    spread_df.columns = names
    return spread_df
```

```python
def normalize(data):
    """
    Normalize every time series in the Data frame
    """
    for i in range(0, data.shape[1]):
        col = data.columns[i]
        data[col] = (data.iloc[:,i]-data.iloc[:,i].mean())/
    data.iloc[:,i].std()
    return data
```

```python
def signal_matrix(data, spread_matrix,n_std=1):
    """
    Return the signal matrix, given dataset and spread_matrix
    """
    n = spread_matrix.shape[0]
    m = spread_matrix.shape[1]
    signal = np.zeros((n,m))
    n_transaction = []

    for i in range(0,len(spread_matrix.columns)):

        t_up = n_std
        t_down = -n_std
```

```

        signal_array = np.zeros(len(signal))

        for j in range(0,len(signal)):

            thresh_close = 0
            sp = spread_matrix.reset_index(drop=True)
            val = sp.iloc[j,i]

            no_positions = np.nansum(signal_array[:j])==0
            long_positions = np.nansum(signal_array[:j])==1
            short_positions = np.nansum(signal_array[:j])==-1

            #OPEN LONG
            if val < t_down and no_positions:
                signal_array[j] = 1

            #OPEN SHORT
            if val > t_up and no_positions:
                signal_array[j] = -1


            #CLOSE LONG
            if val > thresh_close and long_positions:
                signal_array[j] = -1


            #CLOSE SHORT
            if val < thresh_close and short_positions:
                signal_array[j] = 1

            if j == len(signal)-1:
                if long_positions:
                    signal_array[j] = -1
                if short_positions:
                    signal_array[j] = +1

        signal[:,i] = signal_array.T
        s = pd.DataFrame(data=signal)
        s.columns = spread_matrix.columns
        n_transaction = abs(s).sum()
        s.sum()

    return s, n_transaction
```

```
def annualize_rets(r, periods_per_year):
```

```
    """
    Annualizes a set of returns
    """
    compounded_growth = (1+r).prod()
    n_periods = r.shape[0]
    return compounded_growth**(periods_per_year/n_periods)-1
```

```
def annualize_vol(r, periods_per_year):
    """
    Annualizes the vol of a set of returns
    """
    return r.std()*(periods_per_year**0.5)
```

```
def sharpe_ratio(r, riskfree_rate, periods_per_year):
    """
    Computes the annualized sharpe ratio of a set of returns
    """
    # convert the annual riskfree rate to per period
    rf_per_period = (1+riskfree_rate)**(1/periods_per_year)-1
    excess_ret = r - rf_per_period
    ann_ex_ret = annualize_rets(excess_ret, periods_per_year)
    ann_vol = annualize_vol(r, periods_per_year)
    return ann_ex_ret/ann_vol
```

```
def drawdown(return_series: pd.Series):
    """Takes a time series of asset returns.
       returns a DataFrame with columns for
       the wealth index,
       the previous peaks, and
       the percentage drawdown
    """
    wealth_index = 1000*(1+return_series).cumprod()
    previous_peaks = wealth_index.cummax()
    drawdowns = (wealth_index - previous_peaks)/previous_peaks
    return pd.DataFrame({"Wealth": wealth_index,
                         "Previous Peak": previous_peaks,
                         "Drawdown": drawdowns})
```

```
def semideviation(r):
    """
    Returns the semideviation of a return series or Dataframe
    """
    if isinstance(r, pd.Series):
        is_negative = r < 0 #assuming daily returns the mean is approximately 0
```

```
        return r[is_negative].std(ddof=0)
    elif isinstance(r, pd.DataFrame):
        return r.aggregate(semideviation)
    else:
        raise TypeError("Expected r to be a Series or DataFrame")
```

```
def risk_contribution(w,cov):
    """
    Compute the contributions of the assets, given weights and cov.
    """
    total_portfolio_var = portfolio_vol(w,cov)**2
    # Marginal contribution of each constituent
    marginal_contrib = cov@w
    risk_contrib = np.multiply(marginal_contrib,w.T)/total_portfolio_var
    return risk_contrib
```

```
def target_risk_contributions(target_risk, cov):
    """
    Use to find ERC portfolio, risk budgeting function
    """
    n = cov.shape[0]
    init_guess = np.repeat(1/n, n)
    bounds = ((0.0, 1.0),) * n
    # construct the constraints
    weights_sum_to_1 = {'type': 'eq',
                        'fun': lambda weights: np.sum(weights) - 1
    }
    def msd_risk(weights, target_risk, cov): # define the function to create an optimization problem
        """
        Returns the Mean Squared Difference in risk contributions
        between weights and target_risk
        """
        w_contribs = risk_contribution(weights, cov)
        return ((w_contribs-target_risk)**2).sum()

    weights = minimize(msd_risk, init_guess,
                       args=(target_risk, cov), method='SLSQP',
                       options={'disp': False},
                       constraints=(weights_sum_to_1,),
```

```
                              bounds=bounds)
    return weights.x
```

LUISS

Department of Economics and Finance
Equity markets and Alternative Investments

Performance Analysis of Statistical Arbitrage Portfolio

Advisor
Paolo Vitale

Co-advisor
Marco Morelli

716921
Niccolò Fabbri

Academic Year: 2020/2021

Contents

Chapter 1

Introduction

Harry Markowitz described diversification as *"the only free lunch in finance"*, risk can be reduced without sacrificing returns. However, empirical evidence suggests that there is another *"free lunch"*: arbitrages. Arbitrages occur when markets inefficiency generates an opportunity for risk-free profits.

The majority of trades nowadays are executed by algorithms. Institutional investors exploit arbitrages in seconds or even milliseconds, securing profits and restoring market efficiency.

This thesis will cover the topic of Statistical Arbitrage. This approach uses statistical techniques to infer relationships between assets and generate profits in their short-term deviation. However, such a strategy is not risk-free. It is exposed to model risk and security-specific risk. Relationships on which the model is based may be spurious or may not continue in the future due to changes in markets dynamics.

Our analysis wants to investigate if a Statistical Arbitrage trading strategy can take advantage of market inefficiencies to achieve consistent profits. In particular, we will develop an algorithm that backtests such a strategy to analyze its performance in the US stock market. Furthermore, alternative weighting schemes will be tested to assess the importance of asset allocation and risk budgeting in a market-neutral strategy.

Chapter 2

Overview of Pairs Trading

The first practice of statistical pairs trading is attributed to Wall Street quant Nunzio Tartaglia who was working in Morgan Stanley in the mid-1980s[1]. He introduced a trading technique that aims to identify pairs of securities whose prices move together, and if the spread between them widens, an arbitrage opens. However, there is no certainty that stock prices preserve their past behavior in the future; therefore, pairs trading may be unprofitable.

2.1 Cointegration approach

A time series is stationary if the stochastic process generating the series has time-invariant parameters. Stationary times series are easy to forecast, and they are mean-reverting. Empirically we can model a generic stock price time series with a random walk, which is the simplest example of an integrated process $I(1)$. Random walks sequences are unpredictable.

For this reason, the work of Engle and Granger (1987) turned out to be one of the most valuable discoveries both in theory and in practice since many funds adopted a trading strategy based on their model. Statistical Arbitrage based on cointegration analysis relies on the non-stationarity of stock prices.

In particular, time series must be an integrated process of order 1 $I(1)$. This means that the process itself it's not stationary but its first difference is. Two processes $I(1)$ are cointegrated if there's a linear combination of them that is stationary $I(0)$, if two time series are cointegrated they comove. In brief, if y_t and x_t are historical stock prices there are two coefficients α and

[1]Vidyamurthy, 2004, chap 5 p.74

β such that:

$$\alpha x_t + \beta y_t = z_t \text{ where } z_t \text{ is } I(0) \tag{2.1}$$

A long-term equilibrium relationship links these processes. Therefore, the investor should take advantage of the short-term deviations from the equilibrium. The goal is to find cointegrated securities where the spread z_t has a high degree of mean reversion, estimate the cointegration coefficient, and profit on temporary mispricings. Cointegration and its implementation in a trading strategy will be explored in more detail in future chapters since the purpose of this section is to provide the reader with all the necessary tools needed for a proper understanding of the entire thesis.

Chapter 3

Trading Strategy

3.1 Data preparation and pairs selection

The data analysis and the backtesting of the strategy will be performed on Jupyter Notebook with Python. We will use daily data of stock prices of all the components of S&P500 in the period 2014-2021Q2, provided by Bloomberg. Since many companies in the index were included in it during our analysis period, to avoid biases, we will remove every stock that it's been added or removed in that period. This cleaning process will lead us to a DataFrame with 476 stocks rather than 500. To validate the strategy in a forward-looking perspective, we will split the dataset into two parts to analyze the in-sample and the out-of-sample performance. The validation will be performed on a rolling basis. The estimation window will be five years long and the out-of-sample three months. Every quarter the in-sample dataset will slide, taking into account information of the out-of-sample dataset, and the parameter estimates will be updated in a rolling way.

In order to create an active portfolio, we need to select several pairs; the number can range from 15-30 to maximize diversification-transaction cost trade-off. As mentioned in section 2.1 our criterion to identify pairs is Cointegration.

First of all, since Cointegration assumes that the processes are integrated, we have to remove from our basket of securities the stocks which prices form a stationary process. We need to run a unit root test and check whether its parameters change over time. We find out that 13 out of 476 time-series are stationary, so we need to drop them. This process will lead to a cleaned dataset of 463 stocks.

Removed companies, whose stock prices are generated by a stationary process have a total capitalization of 579.989 billions[1]. Since the total capitalization of the S&P500 is around 37 trillions[2], we removed approximately 1.5% of the entire index. In addition, we can see that the companies' betas range from 0.3 to 1.7, underlying that the stationarity does not depend on the company's riskiness.

3.1.1 Cointegration testing and spread series

To find the pairs in our sample, we need to run the Engle-Granger Test on the whole DataFrame. The E-G tests for no-cointegration of a univariate equation and is a two-step process:

1. Determination of the linear relationship.

2. Stationarity testing on the spread series (with ADF test).

The null hypothesis is no cointegration. If the p-value is small, below a critical size, we can reject the hypothesis that there is no cointegrating relationship. Given 463 different stocks, we have 106953 possible pairs. Once we have iterated the statistical test over every possible pair of time series, we find 5306 pairs at a confidence level of 5%. Therefore for each pair, we can construct the spread series z_t:

$$y_t - \beta x_t = z_t \text{ where } z_t \text{ is stationary} \tag{3.1}$$

Where y_t and x_t represent two time-series, and β represents the cointegration coefficient. To compute the latter, we will use OLS, performing a regression of y_t against x_t. The beta of the regression is the quantity of x necessary to hedge 1 unit of y and therefore to maintain the stationarity relationship.

To avoid data snooping bias and select an ideal number of pairs: we will we sort all the 5306 pairs by their half-life[3]. The half-life is defined as the number of periods required for the impulse response to a unit shock to a time series to dissipate by half. It is widely used to quantify the degree of mean reversion.

[1]Data provided by Yahoo-finance!

[2]accordingly to 2021Q3

[3]Since we use daily data, the half-life is expressed in days.

3.2 Trading signals

To backtest the trading strategy, we need a specific rule that, when is triggered, tells us to open or close a position. In our case, we are aware that the spread series is stationary, with a high degree of mean reversion; indeed, we will base our signals on the latter. In particular, we will standardize every spread series to create a level playing field and set up a common rule.

Once the spread series is normalized, its mean will be 0. Since we want to take advantage of the temporary mispricings, we will trade the assets when the spread is above or below a certain threshold because we rely on the fact that it will go back to 0. We want to avoid overfitting, and for this reason, we will set a unique threshold of 1. When the absolute value of the spread series takes a value greater than one, a position on the spread will be opened. Opening a position on the spread means opening a long and a short position simultaneously. Specifically:

- long spread : buy the *Y-asset* and short β on the *X-asset*.
- short spread : short the *Y-asset* and buy β on the *X-asset*.

In particular:

- If the spread takes a value lower than -1, we will open a long position on the spread.
- If the spread takes a value greater than 1, we will short the spread.
- The exit point will be 0.

We want to maximize the performance of our trading strategy. For this reason, in the next chapter, we will examine alternative weighting schemes that may or may not enhance the return or risk profile of the strategy. Finally, we will compare a naive weighting scheme vis a vis risk-based indexation.

Chapter 4

Weighting Schemes

Since our strategy relies on trading 20 different pairs, the topic of budget allocation is crucial. We desire to allocate the budget to the best-performing assets, both in terms of risk and return. This chapter will analyze the capitalization weighted portfolios, with their advantages and disadvantages. Then, it will cover risk budgeting with a particular focus on the risk parity portfolio. Ultimately, it will present the implementation of the equal risk contribution principle in our strategy.

4.1 Capitalization weighted portfolios

In today's day and age, most funds base their portfolio optimizations on Modern Portfolio Theory. In his 1952 paper, Markowitz described portfolio selection with a simple rule: "The investor does (or should) consider expected return a desirable thing and variance of return an undesirable thing.". He stated that there is no optimal portfolio but a set of efficient portfolios. In an efficient scenario, the investor would hold the market portfolio: the portfolio containing every asset weighted by market capitalization.

Cap weighted indexation is the most representative weighting of the market. Moreover, it is the only weighting scheme that is compatible with the efficient market hypothesis. Another feature of the CW portfolios is that they do not change if the market structure remains unchanged. The weight of an asset changes according to its return. They are the most convenient portfolios in terms of rebalancing and transaction costs.

Inefficiency of Cap-weighted portfolios

Despite their popularity and simplicity, CW portfolios tend to be inefficient. They tend to be highly concentrated and poorly diversified, thus exposing the investor to specific risk.

Capitalization weighted portfolios tend to have a second drawback: they have an inefficient exposure to risk factors, regarding the Fama-French three-factor model. It is empirically proven that small-cap stocks outperform large-cap stocks, and value stocks outperform growth stocks. Therefore, holding a cap-weighted means having a positive tilt towards value and large-cap stocks. We are allocating asset inefficiently regarding to SMB Fama-French risk factor.

In addition, CW indexing is a trend-following strategy that incorporates momentum bias and leads to bubble exposure risk[1].

Equally weighted portfolio

A naive approach to diversification suggests creating a portfolio well balanced in terms of dollar allocation. Therefore, the EW portfolio represents the most straightforward portfolio. For a portfolio of n assets, the weight of each asset will be $1/n$. Weights are defined based only on the number of assets, and stocks properties are not considered.

This allocation is ideal if we assume that it is impossible to forecast risk and return. EW portfolio coincides with the efficient portfolio if the returns and volatility are equal and correlation is uniform[2]. In our analysis, we will use the EW scheme as a benchmark, compared with a risk-based weighting scheme.

4.2 Risk budgeting

Risk budgeting is an asset allocation technique that weigh assets basing on their risk. Weights are designed such that each asset contributes in a certain way to the portfolio's overall risk. We will focus on the equal risk contribution approach, where every asset has the same risk contribution to the portfolio's overall risk. This strategy in the industry took the name of Risk Parity.

[1] Roncalli, 2013, chap 3 p.157

[2] Roncalli, 2013, chap. 3 p.164

4.2.1 Risk contribution

A well-balanced portfolio in terms of dollar allocation does not imply well balance of risk. To allocate our budget in terms of risk, we have to define the risk contribution of each asset.

$$RC_i = \frac{\mathsf{w}_i^2\sigma_i^2 + \sum_{j\neq i}^{N} \mathsf{w_i w_j}\sigma_{ij}}{\sigma_p^2} \tag{4.1}$$

Where w are weights, σ_p^2 is the portfolio variance and σ_{ij} represents the covariance between i and j asset.

Equal risk contribution

Risk parity portfolio is analogous to EW portfolio but in terms of risk allocation. The goal of the strategy is to set weights that make risk contributions equal among all assets. The target risk contribution of each security, described in equation 4.1 ,will be $1/n$.

4.3 Implementation

In our review, we consider 20 pairs, therefore 20 time series of returns. We will use the equal risk contribution principle to determine the initial weighting of each pair in the total trading portfolio. In particular, we will compute the return of each pair regarding the in-sample period and then construct the covariance matrix.

Once we have the covariance matrix, we will compute the risk contribution of each asset, and then we will apply the risk parity weighting.

We will set ERC weights, at the beginning of our out-of-sample testing period, 2019Q1. Since the validation period is two years and a half, to make the most of this asset allocation, we will not rebalance the weights. With this weighting scheme, we should have a more balanced portfolio in terms of risk and, therefore, performs better in the backtesting. We will test this allocation, comparing it with the EW portfolio, in the next chapter.

Chapter 5

Performance Analysis

This chapter will analyze the performance of the trading strategy regarding the out-of-sample period (2019Q1-2021Q2). We will backtest the trading strategy with two different weighting schemes presented in the previous chapters. In particular, we will compare the performances of an EW portfolio among pairs, and a Risk parity one. In the first case, each asset's initial budget will be equal, in the second one will be based on the risk contribution.

5.1 Return computation

Since we have quantity constraints, given by the cointegration coefficients, unlike Gatev et al. (2006) analysis, we cannot set the weight of the componetns of the pair to 50% each. Therefore, the weighting of the components of a pair at time t depends on prices. The process of return estimation will be as follows:

1. computes individual return of each asset day by day.
2. weigh such return for the weight of the asset in the pair, given the quantity constraints.
3. weigh the return referred to in point 2 for the weight attributed to that particular pair, in our case: EW and Risk Parity weights.

5.2 Strategy returns

This section will review the strategy performance, provide the returns, risk-adjusted measures, and analyze the strategy's overall riskiness. The strategy performed well in both EW and ERC portfolios. In particular, we generated a net return of 156% over two and a half years in the EW portfolio. The risk parity portfolio generated a net return of 143%. Figure 5.1 depicts the cumulative return of the strategy. Such returns annualized become 43% for

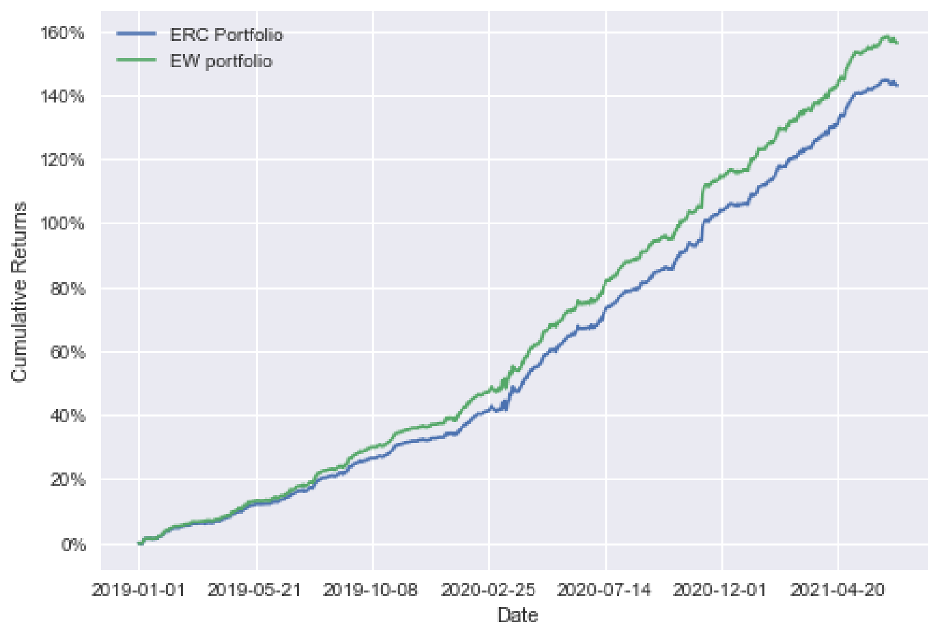

Figure 5.1: Cumulative returns of EW and ERC portfolios

the EW and 40% for the ERC, as can be seen in table 5.1.

	Annualized Return	Annualized Volatility	Maximum Drawdown
EW	0.438599	0.050172	-0.020758
ERC	0.409259	0.047062	-0.021270

Table 5.1: Summary statistics of portfolio returns

5.2.1 Pairs characteristics

In this section, we will analyze how pairs contributed to the portfolio return. The best performing pair has been *Servicenow - Mastercard*. It generated a return of 343% over his initial budget. On the other hand, the worst pair in terms of returns has been *American Electric Power -Eversource Energy*, with a 33%. Hence, the former has outperformed the latter by ten times. This confirms the robustness of the analysis; pairs are keeping their behavior out-of-sample. In particular *Servicenow-Mastercard* was the second-best performing pair, and *American Electric Power-Eversource* was the worst-performing one, regarding the in-sample analysis.

Since the number of trade is approximately the same, 64 and 58, each transaction on the former secured a much higher profit than the latter.

Regarding the EW portfolio, pairs that contribute the most to the overall return are the pairs that performed better. However, this is not true in the ERC portfolio since weights are based on in-sample risk contribution. Hence, even though a pair generates an abnormal return, it can be weighted less than $1/n$[1] due to its risk contribution. In particular, in our case, the most performing pairs are underweighted, apart from one specific pair: *UnitedHealth Group - CDW Corporation*. This pair is optimal, has a low risk contribution, and generates an out-of-sample return of 292%. Applying the ERC, its weight on the portfolio rises; therefore, its contribution to the portfolio return in the Risk parity portfolio is higher than in the EW. However, this is not sufficient to balance the effects of the weight reduction in high performing assets.

5.2.2 Maximum Drawdown

An essential indicator of the downside risk is maximum drawdown. It is described as the maximum observed loss from a peak to a trough of a portfolio.

Risky portfolios tend to have higher drawdowns than "non-risky" portfolios. In our case, EW represents the naive diversification in terms of risk, and therefore, the risky portfolio. ERC, on the other hand, represents the "conservative" portfolio.

However, we have an unexpected outcome. The maximum drawdown is higher in the ERC portfolio, -2.12% compared to -2.07%. Even though they differ only by five basis points, this result proved that the "conservative"

[1] most of the times

weighting scheme suffered market turbulence. Indeed maximum drawdowns in both portfolios manifested the same day, March 17, 2020. That day financial markets were in the middle of the covid crisis. In particular, we can find an analogy between the performance of S&P500 and our strategy. The maximum drawdown of S&P500 occurred on March 23, 2020, with a -34%.

Furthermore, if we look at the VIX index, also known as the *fear index*, which represents the stock market's expectation of volatility based on S&P 500 index options, we can see that on March 16, it surged around 43%, closing to a record high of 82.69, surpassing for the first time the peak level of November 21, 2008.

Our strategy aims to market neutrality. However, evidence suggests that there is a minimum relationship between S&P500 and our portfolio returns.

Strategy's Beta

We want to analyze the relationship, between S&P500 and the returns of our strategy, that came out of drawdown measures. In particular, we want to check if the beta of our strategy is negligible. Therefore, we will perform a regression between the market excess returns and the excess return of our portfolios. Results of the regression are displayed in table 5.2 and 5.3.

We notice that the betas of both portfolios are close to 0. In particular, given p-values, we see that both betas are not significant. Results show market neutrality. In EW and ERC portfolios, beta cannot explain returns, proving they are uncorrelated with the market. At the same time, Jensen's alpha is positive and significant. They are 0.031 and 0.029, respectively. Such high alphas state that the unexplained returns are approximately 3,1% and 2.9% monthly. Which annualized represents the total return of the strategy. Outcomes have shown that the strategy is indeed market neutral.

EW	Coefficient	Std. Err.	P-val.
Beta	0.0128	0.045	0.780
Alpha	0.0315	0.003	0.000

Table 5.2: EW portfolio regression

ERC	Coefficient	Std. Err.	P-val.
Beta	0.0304	0.044	0.494
Alpha	0.0293	0.003	0.000

Table 5.3: ERC portfolio regression

5.2.3 Risk-Adjusted Measure

It is well known that exists a trade-off between risk and return. If we evaluate funds' performance by just looking at returns, we make a mistake. We need to correct returns with a risk measure. The most famous ratio in such a sense is the Sharpe ratio.

We can see from figure 5.1 that the returns are pretty similar. Indeed annualized volatilities, as can be seen in table 5.1, are 5% and 4.7%, respectively. Thus, we give up a 3% percent annual return to gain a -0.3% portfolio volatility. Given such low volatilities for both portfolios, it's not worth renouncing on a 13% return to save 1% of riskiness. This, as can be seen in table 5.4, is confirmed by the Sharpe ratio, which is higher in the EW portfolio rather than the ERC.

	EW portfolio	ERC portfolio	S&P500 index
Sharpe Ratios	8.720457	8.673775	0.855131.

Table 5.4: Sharpe ratios of different portfolios

The risk parity scheme should have outperformed the naive allocation in terms of Sharpe ratio but reduced return more than reducing risk. In other words, results have stated that the ERC portfolio is inefficient.

On the other hand, Sharpe ratios are huge compared to market Sharpe ratios. In particular, the Sharpe ratio of the strategy is approximately ten times bigger than the one in the market. We cannot compare a passive buy-hold strategy with an active one. Even though we compare it with the active management industry Sharpe ratios, a value of 8 is really high. These superior performances will be justified in the last section.

5.2.4 EW vs. ERC

The strategy has been backtested with two different weighting schemes. Naive diversification, where every pair weights $1/n$, and Risk parity, where pairs are weighted to achieve equal risk contribution. The ERC portfolio, in our strategy, turned out to be inefficient. It produced worse results than the EW portfolio in terms of returns, drawdown, and risk-adjusted return.

Our analysis infers that ERC is not optimal to implement in a statistical arbitrage trading strategy. Hence, given the low risk of the strategy itself, if an investor has to allocate his budget among pairs, results suggest that creating an equally weighted portfolio is a much better solution.

5.2.5 All that glitters is not gold

This section will analyze the issues and the hidden costs of the trading strategy. First of all, our analysis does not consider transaction costs. Since we invest in an active strategy, there will be a lot of transactions. In particular, we executed a total of 1266 trades on the pairs over two years and a half. Gatev et al. (2006), states that transaction costs are around 50 basis points. If we compute a naive estimation of transaction costs, we can arrive at a total cost of 30%. Leading annualized return to 32% for the EW and 29% for the ERC portfolio. In addition, there are hidden costs that we are not taking into account. Most of them are related to short-selling. In particular, lending fees, stop-loss margins, and margin calls.

Lending fees can be daily; for this reason, positions opened for too long may be less profitable.

Margin calls occur when the value of an investor's margin account falls below the broker's required amount. As a result, the investor is obliged to post more collateral or close the position, reducing the strategy's profitability again.

The validation period covers two years and a half. So even though there is a financial crisis in the middle of it that tests the robustness of our market neutral strategy over different market conditions, the period may not be long enough to assert the absolute profitability of the strategy.

However, the performance of the strategy was excellent. For this reason, taking into account these costs will inevitably reduce the returns but will not invalidate the results.

Chapter 6

Conclusions

In this thesis, we introduced the concepts of Pairs trading and long-short equity strategies. We reviewed stationarity, cointegration, and market neutrality. In addition, we presented the commonest weighting schemes, emphasizing their benefits and drawbacks, comparing them with risk budgeting techniques.

As thesis core, we developed an algorithm that backtests a statistical arbitrage trading strategy. In particular, it tests for cointegration relationships among a broad set of assets and selects the most tradable pairs. Subsequently, it creates trading signals to enter or exit a trade. Finally, it computes the daily P&L and summary statistics of the portfolio.

Furthermore, we proposed a new weighting scheme for a statistical arbitrage portfolio that relies on the equal risk contribution principle. Such allocation aims to enhance the risk-adjusted performance of the portfolio, reducing its overall risk.

The algorithm tested the strategy in the US market regarding the period 2014-2021, with two years and a half of out-sample backtesting. The trading strategy generated an abnormal return of 40% annualized, with extremely low drawdowns and high risk-adjusted measures. These outcomes have to be seen through the lens proposed in section 5.2.5. In addition, results showed that our returns are uncorrelated with the market. Therefore, the strategy has been successful in achieving market neutrality.

However, the risk parity weighting scheme did not achieve its purpose. The equally weighted portfolio outperformed the equal risk contribution portfolio in terms of returns, maximum drawdown, and Sharpe ratio. We conclude that given the low risk of the strategy, a more balanced allocation in terms of risk generates inefficiencies. Results suggest that we are exposed to upside risk more than downside risk. For this reason, balancing assets inversely proportional to their risk contribution ends up reducing returns rather than limiting risk.

In conclusion, within our analysis, the strategy performed well. Given such promising results, this strategy will be improved and tested in the real market.

Bibliography

S. Benartzi and R. H. Thaler. Naive diversification strategies in defined contribution saving plans. *American economic review*, 91(1):79–98, 2001.

Z. Bodie et al. *Investments*. Tata McGraw-Hill Education, 2009.

M. Clare and Thomas. An evaluation of alternative equity indices part 1: Heuristic and optimised weighting schemes. 2013.

E. J. Elton and M. J. Gruber. Estimating the dependence structure of share prices–implications for portfolio selection. *The Journal of Finance*, 28(5): 1203–1232, 1973.

R. F. Engle and C. W. Granger. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, pages 251–276, 1987.

E. Gatev, W. N. Goetzmann, and K. G. Rouwenhorst. Pairs trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies*, 19(3):797–827, 2006.

R. A. Haugen and N. L. Baker. The efficient market inefficiency of capitalization–weighted stock portfolios. *The journal of portfolio management*, 17(3):35–40, 1991.

F.-S. Lhabitant. *Handbook of hedge funds*, volume 332. John Wiley & Sons, 2007.

T. Roncalli. *Introduction to risk parity and budgeting*. CRC Press, 2013.

G. Vidyamurthy. *Pairs Trading: quantitative methods and analysis*, volume 217. John Wiley & Sons, 2004.