LUISS

Double Degree Master's Thesis

Department of Economics and Finance
Major in Finance

Chair of Asset Pricing

Multifactor Risk Analysis in European Equity Mutual Fund Industry

Prof. Borri Nicola
SUPERVISOR

Ligabue Filippo
CANDIDATE
ID: 734261

Prof. Santucci De Magistris Paolo
CO-SUPERVISOR

Academic Year 2021/2022

Table of Contents

Introduction

This thesis is an expansion of my term paper called: "Multifactor Analysis and Mutual Fund Application". The topic of my research is structured in two dimension whose are partially dependent: the first one is the multifactor part and the second one is the downside risk part.

The first goal of my thesis is to verify empirically how much the factor investing is important in mutual fund field and in particular in the strategies of the fund managers of the European Equity Mutual fund sector. After summarising the theories regarding the asset pricing factors, in order to achieve that I developed some Python code to numerically verify some hypothesis and have a clear result of how is relevant the factor investing theory in a significant sample of European Equity Mutual funds. For the literature review I started from the classic French Fama factors then and I decided to analyse also the more recent factors of Frazzini and Pedersen whose are punctually documented in the AQR Capital Management database.

The second aim of my thesis consist in analysing the downside risk phenomena. This result was obtained due to a deep study of the most relevant theories that were written regarding this specific topic, in particular statistical indicator like the downside risk and the coskewness.

Then I did some computations in Python in order to attest if the downside risk is taken into account in my sample, which I constructed in order to be representative of the market segment whose is my focus.

I focused on this study because I was really interested in the asset pricing after attending the university course and in particular, I was impressed when we did some lessons regarding these topics. So, I decided to do this research in order to deepen my personal knowledge of this arguments and build mathematical tests and models to test hypotheses.

I hope that my contribution to the literature review could be important since there is not a large number of research who implement these theories and verify empirically the reply of the practitioners to the factor investing strategies and if mutual funds managers consider downside risk in their portfolio assets allocations.

Chapter 1 - Literature review: explanation of each factor

For the multifactor analysis part, I decided to focus my research in seven factors whose are:

- MKT, is the market factor, it is the first factor of the French-Fama three factors model which capture the excess return of the market portfolio over the risk-free rate.
- SMB, is the small minus big factor, it is the second factor of the French-Fama three factor model, it represents the excess return made on size effect.
- HML: high minus low, is the third factor of the French-Fama three factor model, it represents the value-growth factor.
- UMD: is the momentum factor proposed by Mark M. Carhart in 1997, and successively studied and modified, in my research for the calculus I consider the formula used by of Asness, Frazzini and Pedersen in their 2013 research.
- QMJ: it means quality minus junk; it represents the factor that explains the trend to choose for quality stocks.
- HML D: high minus low devil, it is a revisitation of the classic HML factor by Asness and Frazzini in 2013.
- BAB: betting against beta, is one of the most recent factors developed by Asness and Frazzini in 2014, which represent the empirical evidence that lower beta stocks had greater returns over time.

I choose these factors because the literature review and the empirical evidence on this topic highlight that:

- investing with strategies that are focused on these factors in average generate time-varying expected risk premium,
- there is an important correlation between the returns over time of this strategies of stock picking and real activity,
- the selected factors are the more representative of all current asset classes and they are present not only in stock markets but also in the others.

The founding fathers of the multifactor analysis are Kennet R. French and Eugene F. Fama, the second one won a Nobel Memorial Prize in Economic Sciences for his empirical analysis of asset prices. In particular, they discovered three common risk factors that explain the cross section of average returns on Amex, NASDAQ and NYSE stocks for the time period from 1963 to 1990 in respect to monthly returns.

Later, it was empirically demonstrated that has a valid explanation in almost all international stock markets. The result of this important study was published in 1992 in the paper: "Common risk factors in the returns on stocks and bonds" and it takes the name French-Fama three factor model.

Even if in their research they made an important implementation also in the bond field, the main focus of my term paper is the equity market. The method that was used to explain these relationships was the time series approach of Black, Jensen and Sholes of 1972. The regression used as dependent variable the excess stocks return over risk free rate and as independent ones the variables of economics interest, in particular portfolios sorted on the factors characteristics of interest. The reasons of that approach were mainly two: the first one was that this regression intercepts provides an evident explanation of the relationship between the variables and in particular of how good the combinations of the factors capture the cross section of average returns. If this type of regressions returns alphas that are statistically near zero, we could say that our factors of interest verify my hypothesis of their explanatory power. Secondarily, the ratio behind pricing assets implies that book to market equity, size and other measures that are correlated to returns need to reflect sensitivities to common, shared, and therefore non-diversifiable, risk factors in returns.

The practical implication of time series regression is an important instrument to highlight this issue, most of all numerical results like R square and slopes of the regressions shows us how much of the variation in stock returns is captured by the risk factor of analysis, in this context book to market equity and size.

For the stock portfolios the economists studied, the three-factor regression portion including excess market return and size-simulated return and the BE/ME factor was close to 0. Therefore, the market factor and the proxy size and book-to-market value for the risk factor in the correlation than seem to explain the cross-section of average stocks returns well.

As I wrote before the three model-factors are: MKT, SMB and HML. The first one is representative of the average market return; this factor is constructed as the difference between the excess market portfolio return and the one-month relative T-bill rate (in formulas Rm-Rf).

In order to obtain the market return (Rm), we have to construct six portfolios based on a combination of size and Book equity/Market equity, so practically the economists have divided all the markets stocks in two sets based on these two characteristics and then into more sub-categories. In particular, when we deal with size the stocks are divided in two samples based on their relative market capitalization, which is number of shares outstanding of a firm time the relative market price, the classification is made by the median of our sample.
The result is represented by two sub-portfolios called Small (S) or Big (B). The median of NYSE stocks on CRSP is computed every June of each year of the sample from 1963 to 1991, as the previous explained method and then this division it is used to divided also the Amex and the NASDAQ (after 1972) markets.
In sorting for Book to market equity value the economists created three sub-sets by the following subdivision ordered in increasing trend: first 30% of stocks were denoted by Low (L), intermediate 40% named Medium(M), highest 30% called High (H).
For computing the market price French-Fama method consider the 30th of June price and for computing book value the 31st of December data, this was extremely discussed by Asness and Frazzini in 2013 when they introduced their modified version of HML factor named HML Devil. However, this aspect will be discussed later.

The difference between the two division and the more categories that we have in BE/ME analysis is made by the evidence that book to market equity has a strong impact in average stock returns than size effect. In the final passage there was a crossroad between all the subdivisions to form six portfolios S/L, S/M, S/H, B/L, B/M, B/H, these comes from the combinations of our 5-division made by dimension and book to market ratio.
Coming back to the market risk factor, Rm (market return) it is computed as the sum between the return on negative book to equity stocks that are excluded from the portfolio and the return of six book to market equity- size value weighted portfolios.

SMB, small minus big, is the second factor of analysis and it reflects the size effect, it came from the empirical evidence that small firms in average had a smaller profitability compared to the biggest ones. This factor is computed isolating the influence of BE/ME. In other words, it describes the negative relation between average returns and size.
It is computed as the difference between the arithmetic average of the small stocks' portfolios (so in practice Small/Low, Small/Medium, Small/High) and the arithmetic average of the big stock's portfolio (Big/Low, Big/Intermediate, Big/High).

HML, high minus low, is the last factor of the model and it represents the so-called value premium and it empirically represents the evidence that over time value stocks present better results than growth stocks. The first group is represented by firms who have a high BE/ME, ratio so the stock price is relatively low compared to the stock book value; they present smaller earnings on assets, this is numerically showed by French and Fama that consist of five years prior to book to market ratios measures and five years before. Growth stocks in oppositive have a high market price relative to firm book value and they highlight in average high earnings.
So, in opposite from size effect the value effect it is caused by a positive relation between average return and book to market ratio.
The factor in opposite to SMB is computed as the difference between the arithmetic average of the big book to equity portfolios (so in practice Small/High, Big/High) and the arithmetic average of the smallest book to equity portfolios (Small/Low, Big/Low).

The UMD factor is the momentum factor which analyses and documents the persistence that is present in the mutual fund performance. It was primally documented in interesting research by Carhart in 1997 in his paper: "On persistence in Mutual Fund Performance".
In this study was introduced one of the first version of the momentum factor which was called PR1YRt. The economist added this new factor on the previous French-Fama three factor model to implement a new four-factor model in 1995. This new additional independent variable in the regression replicate the Jegadeesh and Titman's reported one year momentum anomaly (discovered in 1993). This variable is a factor mimicking portfolio for one year return momentum.

This new dimension of analysis is the result based on the evidence of a lot of studies that prove a relative slightly consistence in the subsequent monthly return in a past year fund which had returns higher than market average. It was empirically demonstrated that the returns of the top performer last year fund were very high correlated with this new factor, that we call one-year momentum factor, in contrast the worst performers demonstrated a negative relationship with that.
Carhart took a sample of returns of stocks from Nasdaq, Amex and NYSE from July 1963 to December 1993 and he did the following regression, taking all the measures in monthly data returns:
(Rp-rf) =alpha+b1RMRFt+b2SMBt+b3HMLt+b4PR1YR

The computations follow the French-Fama method, RMRF stands for MKT factor.

These are the results:

Factor Portfolio	Monthly Excess Return	Std Dev	t-stat for Mean = 0	Cross-Correlations VWRF	RMRF	SMB	HML	PR1YR
VWRF	0.44	4.39	1.93	1.00				
RMRF	0.47	4.43	2.01	1.00	1.00			
SMB	0.29	2.89	1.89	0.35	0.32	1.00		
HML	0.46	2.59	3.42	−0.36	−0.37	0.10	1.00	
PR1YR	0.82	3.49	4.46	0.01	0.01	−0.29	−0.16	1.00

Table 1: From Carhart, Mark M. "On persistence in mutual fund performance." The Journal of finance 52.1 (1997): 57- 82, page 62.

The results show a notable reducibility in the average pricing errors relative to previous models, the mean absolute value of the Carhart four factor model is less than half than French-Fama three factor model (0.14 % versus 0.31%). The table shows that the cross-correlation between factors is statistically not so relevant so what could be said in this particular case is that he do not consider multicollinearity problems in the process of the estimation of the relative loadings of the model.

To summarise, the research demonstrates that buying last year's winners is a viable strategy to take advantage of the full-year momentum effect of Jegadeesh and Titman (1993) with virtually no transaction costs, as real transaction costs are passed on to long-term mutual fund holders. However, the strategy of selling shares at net asset value should not be considered as a stable because it is universally used. The long-time equilibrium to compensate mutual fund for their damage impact on their performance requires them to charge mandatory transaction fees both for the outgoing and incoming investors. In conclusion, this study demonstrates that investors must focus their attention to last year founds winner because it is shown that, with a great probability, they will continue to have good result next year (but not straightforward). Furthermore, it was shown that all the cost relative to an investment, in particular the transaction costs, load fees and expense ratios negatively affect the return on investment.
In my thesis I used the Asness, Frazzini, Pedersen interpretation of momentum factor, which is defined as UMD that means Up Minus Down. This construction of this factor is based on size and last twelve-month performance of the sample portfolios. The sorting by size follows the French-Fama procedure, in particular in every month at the end there is the size sampling

made by market capitalization criterion. Moreover, when they considered the international stocks, they did not do the median instead they took as breakpoint the relative 80th percentile in each country. Then the Up Minus Down factor is computed as the following formula: Up Minus Down=0.5(Smallest portfolio with highest returns+ Biggest portfolio with highest return)- 0.5(Smallest portfolio with lowest returns+ Biggest portfolio with lowest returns).

QMJ, or quality minus junk, is the factor that represents the trend that characterizes the investment idea of going long on stocks that are considered as quality and go short on stocks that are considered junk. The overloading on this factor was demonstrated that earned positive risk adjusted returns in the global markets (more in detail in the USA and over 23 developed countries). The developed countries considered are the ones that are represented in MSCI World Developed Index on the 31st of December 2012.

The logic is demonstrated by empirical evidence thought years, Asness-Frazzini-Perdersen studied this phenomenon in 2013 and they made an important research to prove that. They started from their definition of quality security: "Quality security as one that has characteristics that, all-else-equal, an investor should be willing to pay for a higher price for: stocks that are safe, profitable, growing, and well managed".[1]

The price of quality is not constant over time, while it fluctuates especially during bubbles and market crisis moments, it was discovered that if quality is cheap, we will have probably a future positive high return on Quality Minus Junk factor. In order to compute and identify the quality characteristics into a stock, we need to distinguish and highlight the elements which lead to a higher price than the average of the market (or comparable stocks). Asness, Frazzini and Pedersen took as reference for this purpose the Gordon growth model, in particular they deal with price book value ratio because is more stable in the cross section and over time. The model defines price to book value as: profitability times payout ratio over required return minus growth rate. The next step of the procedure is to analyse each variable of this ratio and search for measures that can represent it. It was delineated that there are four fundamental characteristics that we have to highlight when we deal with quality stocks. The first one is safety, because in ceteris paribus conditions investors are willing to pay higher stocks which have a lower required return than the market average because they are commonly considered as safer assets. The second one is growth, which is considered as the future possibility of growth

[1] From Asness, Clifford. S., Andrea Frazzini and Lasse H. Pedersen, 2013, Quality Minus Junk. Working Paper, New York University, page 1.

of a firm, principally based on the increasing trend in profits. When we deal with that characteristic of a firm the past time horizon of growing is stabilized in prior five years from today. The third one is profitability, which it was delineated numerically as the profits for each unit of book value the relative firm stock price.

The last one is the payout which can be described as the percentage of earnings paid to shareholders. We should consider the nature of payout because the earnings are reinvested in a good way into the firm so that can generate a sustainable and consistent growing, instead of being given to investors. In other words, if a higher payout is associated with lower future profitability or growth, then this shouldn't call for a higher price, but if we hold all other factors constant, a higher payout should be considered to have a positive relation with price.

In order for the market to price these properties reasonably, they must measure and predict future properties in advance, i.e. they must be durable.

The economists have shown that this is indeed the case; stocks that are profitable, growth, safety and high paying will continue to exhibit these characteristics on average over the next five or ten years.

The construction of QMJ portfolio took the basis from French-Fama model, so in practice the portfolio construction starts from the recognition of quality that the economist measures as a z score based on the four important pillars, which are: payout, growth, safety and profitability. When they had a full rank of the portfolios, they took the French Fama method classification of size, then they compute the Quality Minus Junk factor return as the difference between two elements. The first one is represented by the half of the sum between the smallest quality portfolio and the biggest quality portfolio. The second one is made by the half of the sum between the smallest low-quality portfolio and the biggest low-quality portfolio.

To do this work they considered a large statistical sample made by 39308 stocks, from January 1986 to December 2012 and they considered their relative monthly returns . As the reference of risk-free rate for compute the excess return for each stock they considered the U.S. monthly Treasury bill rate . They follow a standard convention and align the variables to the company's fiscal year end, which ends in calendar year t-1 through June of calendar year t.

The portfolios based on quality characteristic exhibits negative market, value, and size exposures, positive alpha, relatively low residual risk. During market downturns the QMJ factors in average present a good return, this because we could notice a risk-based interpretations based on covariance with market crises. Rather than exhibiting downside risk, QMJ tends to exhibit mild positive convexity, meaning it benefits from flight quality during crises.

The HMLD or High Minus Low Devil is a modification of HMD factor based on Asness and Frazzini definition of Book to market in their 2013 research. They proved that the French Fama approach to compute that factor it is not so precise but contextualized in that era it was partially correct, because the momentum factor was not explicitly studied and documented well in the literature yet. In the classic HML in the computation of book to price ratio for each stock of the portfolio there are used book price and market price which are not updated and they are considered as old because between the two value there was always present a temporal discrepancy which could range from six to eighteen months. This because there was present the convention that the market price value is took on 30th of June and the book price on the balance sheet year end so on 31 December. This primary explanation behind this choice made by French and Fama was the security of having book value data at the time when the portfolio is rebalanced or initially formed. Secondarily for a similar reason they known that 30th June price is available on the 30th of June rebalanced date. Asness and Frazzini, with their High Minus Low Devil modification, demonstrated that using more updated prices can lead to a relevant improvement in the portfolio combined strategy and especially elucidate the dynamic relationship between value and momentum strategies.

To define it formally, the subscript devil indicates that to compute book to market ratios we scale book equity (BE) by the current total market value of equity (ME) at the end of each month following Asness and Frazzini (2013). There is considered a monthly updating of size and book to market breakpoints in order to maintain a value weight sorting in the portfolio.

In mathematical term, the book value per share in month t, is measured as:

log(Modified book price/Market price at t). The modified book price is equal to:

Book price*Cumulative adjustment factor at time t/ Cumulative adjustment factor at the fiscal yearend.

Remember then this standardization by logarithm must be applied also when we deal with excess return so when we subtract the risk-free rate, we consider it as log(1+Rt,s) where t,s indicates the time period when this risk-free rate is applied.

The two-economist demonstrated that if we consider the current book to market ratios in year or monthly terms, the value measure shows better results than French-Fama method.

One interesting research that validated their correlation between value and momentum strategy, that was also an important motivation behind the Fama-French factor interpretation, was realized in Japan market from 1988 to 2011.

The table below represent the results of a regression of UMD factor returns on MKT, SMB, STR and the HML interpretations. In particular the classic High Minus Low factor is the lagged one. Each column represents the results with different High Minus Low factors, we could notice that the negative correlations between value and momentum in last column, and if we adjust for that we could see the incredible value added of this strategy in Japan market in that period.

	(1)	(2)	(3)	(4)
Alpha	1.15	2.26	**7.34**	**12.04**
	(0.34)	(0.68)	(2.37)	(4.49)
MKT	**−0.23**	**−0.16**	**−0.27**	**−0.24**
	(−5.34)	(−3.38)	(−6.41)	(−7.13)
SMB		**−0.19**	−0.05	0.03
		(−2.54)	(−0.63)	(0.42)
STR		**−0.32**	**−0.31**	**−0.18**
		(−4.73)	(−4.95)	(−3.34)
$HML^{annual,Lagged}$		0.14		
		(1.21)		
$HML^{annual,current}$			**−0.58**	
			(−6.60)	
$HML^{monthly,current}$				**−0.80**
				(−12.66)
R2	0.09	0.19	0.30	0.49

Table 2: From Asness, Clifford, and Andrea Frazzini. "The devil in HML's details." The Journal of Portfolio Management 39.4 (2013): 49-68, Exhibit 7.

BAB factor, means betting against beta, represents the empirical evidence that the investors could produce important risk adjusted returns if they take a long position with a lot of leverage in low beta assets while they go short on high beta assets. This factor was presented by Andrea Frazzini and Lasse Heje Pedersen in 2013.

The pillar behind this strategy is that many investors have leverage and margin constraints, so that type of investors typically bid up high beta assets and those assets are commonly associated with a relative low alpha. This attuite of selecting assets with highest beta is because lower beta assets require more leverage. This explanation comes from the basic principle of the Capital Asset Pricing Model that showed that investors want the higher sharp ratio in their portfolio and they use the leverage to obtain that, this obviously depends on the risk attitude and the borrowing capacity of the subject.

In order to demonstrate the nature and the relevance of Betting Against Beta factor the two economists constructed a dynamic model based on leverage constraint. This included a large sample of assets based on 20 international stock markets, future markets, T-bill markets and credit markets. This model tended to represent as more possible investors as possible with and without leverage and margin constraints. The reason behind the construction of this model was to fully show, report and explain the features behind the betting against beta factor and the loading on this factor in the investing style.

All the procedure and the factor willingness are based on five important prepositions, for each one of them it was demonstrated their empirically implementation by testing their relative statements in a broad sample of data.

For constructing the Betting Against Beta factor, they considered a substantial set of asset classes, including stocks, corporate bonds, commodities, credit indices, foreign exchange, country bonds, equity indices and US treasury bonds. The sample covers different period all ending in 2012, for the majority starting from 1989, while the US stocks considered with CRSP value weighted index are took in consideration from 1926. All the excess returns that they took from calculations are above the one-month T-bill rate.

The process of constructing each betting against factor started from ranking on its estimated beta of each security in ascending order. Then there is a subdivision in two big portfolios, the low beta and the high beta, the distinction of the two classes in made by the asset class median or if we deal with international equity the country median. Then the process is carried out on an ongoing basis every month with a continuative rebalancing.

The authors defined the BAB portfolio as: "The BAB is the self-financing zero-beta portfolio that is long the low- beta portfolio and that shortsells the high-beta portfolio."[2]

The return of a BAB factor in time t+1 is computed as following:

$$return\ BAB_{t+1} = \frac{1}{\beta low_t}(return\_low_{t+1} - risk\_free) - \frac{1}{\beta high_t}(return\_high_{t+1} - risk\_free)$$

The model which has led Frazzini and Pedersen to demonstrate the existence and effectiveness of the factor is based on some predictions and pillars as mentioned before.

The essential finding that underlies the Betting Against Beta is that high beta of a stock is

[2] From Frazzini, Andrea, and Lasse Heje Pedersen. "Betting against beta." Journal of Financial Economics 111.1 (2014): 1-25, page 9.

associated with a low alpha. The investors' goal is to maximize the Sharpe ratio and for doing that they must search for security with a beta less than one.
Alpha of a security can be shown that is equal to one minus the relative beta of the security times a Lagrange multiplier that control for the funding constraint tightness. This is because the model created by the economist have a market line which is quite flat, similar to Black ones of 1972, and the slope of that line is directly proportional to the funding constraints. Secondly, it was endorsed that in the investments there is a sort of beta compression: in concrete terms when finding liquidity risk have an elevate risk there is a compression of securities betas in the cross section to one.

The model implies an important relation: the expected return of Betting against beta factors, $E_t\left(r_{BAB_{t+1}}\right)$, are positive and are related with the Lagrange multiplier of fund tightness (denoted by ψ_t) and with a ratio that represent the spread between the highest and lowest beta portfolios, this is expressed by the following formula:

$$E_t\left(r_{BAB_{t+1}}\right) = \frac{\beta high_t - \beta low_t}{\beta low_t \beta high_t}\ \psi_t$$

Finally, a direct consequence of some previous relations is that when it is taken in consideration a bad time characterized by a tighten funding liquidity constraint, we can expect that the betting against beta has a negative return while it is expected an increasing trend in the future returns.

Chapter 2 – Downside risk theory

In my thesis I will treat the phenomena of "Downside Risk" and how the investors take that in consideration when they allocate their wealth into assets who invest into.

The explanation behind that derives from the fact that is common sense to be sensitive and care more about downside losses instead of upside gains. However, if an agent is exposed to "downside risk" and the market is currently in a declining period that agent is going to be loss more than the average investor because his portfolio is characterized by assets that covary strongly with the market.

For being in that position that agent must require a sort of gain during other market periods to cover him from the possible future losses, which is commonly denoted as downside risk premium. This is due to the fact that there are present some assets which move more downward in declining market than upward in a rising market that stocks are the explanation behind the downside risk phenomena.

These investments statistically have high average returns, this can also be seen as a risk-return trade off and if we follow the previous logic behind that the assets relative downside beta can be seen as a risk attribute.

In the literature there are present some economic indicators which are directly related to downside risk. For example: the downside and upside betas, coskewness and cokurtosis.

These not directly measure "the risk premium" but they are related to the directly consequence of the strong relation that I underwrite before, which relate the high average return of the assets that present a strong covariation with the market when it is in a declining period. All these measures demonstrated that there should exist a sort of downside risk premium but they not directly computed it.

In particular, regarding the citated estimators some authors underlined the importance of the coskewness (or Conditional Skewness) as it is empirically demonstrated in a lot of studies that investors are looking for stocks with high coskewness and assets that have a relative low value of that measures will tend to have high average returns. This important observation has been undelighted in 2000 by Harvey and Siddique in a relevant study named: "Conditional Skewness in Asset pricing Tests".

The central focus of the study is the empirical demonstration of their theory, which states economists can confirm that expected returns should include rewards for accepting this risk when asset returns have systematic skewness. Indeed, this intuition is formalized with an asset

pricing model that incorporates conditional skewness. The outcome proved by their results is that conditional skewness helps explain the cross-sectional variation of expected returns across assets and it is relevant even when factors based on size and book-to-market are included. Hervey and Siddique (2000) predict that lower coskewness should be associated with higher expected returns. The importance of systematic skewness is established in the economy and it commands a risk premium, on average for their statistical sample, of 3.60 percent per year.
What is deductible by their important result is the suggestion that the momentum effect is related to systematic skewness. They proved that the low expected return momentum portfolios have higher skewness than high expected return portfolios. This is an interest study because it matches the momentum factor with coskewness and studied an important correlation between them.
Downside risk is not the same as coskewness risk, because downside beta clearly means that the market is going down in a non-linear fashion, whereas co-skewness statistics do not explicitly highlight the asymmetry between falling and rising markets, even when co-skewness may. The same is true over time. Since skewness reflects some aspects of downside covariation, we take particular care to control for congruence risk when assessing downside beta premia.

One of the first measure that was discussed to control for downside risk is the downside beta, this measure was introduced by Bawa and Lindendenberg in 1977.
The concept took the basis from the beta formula of the classic Capital Asset Pricing Model (CAPM), but with the important distinction that the beta of the asset is divided into two measures: the downside beta and the upside beta.
The criterion decided by the economists was to divide the beta and analyse their relative effects when the market returns are higher or lower than the average of them.
Considering the subject of my research the most relevant part should be the one represented by the assets who are represented by a downside beta.
However, as said before, it is statistical and economically relevant to study the direct distinction between these two indicators in their influence on returns between them.
By construction the classic Capital Asset Pricing Model beta, and these two betas present a level of dependence from each other.
The downside beta is computed as:

$$\beta^{-} = \frac{cov(r_a, r_m | r_m < \mu_m)}{var(r_a, r_m | r_m < \mu_m)}$$

Were, r_a stands for the return on the asset of interest, when I deal with asset return as a standard convention, I took the returns in excess of the conditionally risk-free rate (return). r_m referred to the return on the market and μ_m on the average return of the market.
By opposite the upside beta is defined as:

$$\beta^+ = \frac{cov(r_a, r_m | r_m > \mu_m)}{var(r_a, r_m | r_m > \mu_m)}$$

The components of the formula are the same of the previous one except in that case obviously are regarding the upper part of the average market return distribution.
The explanation behind the introduction of an upside beta took the basis of the concept of an investor with a Disappointment Utility, which consider all the specifications in our context can be realistic. With that the preference function the individual is willing to accept a reduction in his return for a stock that it is consider having a sort of upside potential instead of the average, this could be seen as a sort of discount. At ceteribus paribus conditions, if we consider two stocks that have the same downside beta but one has a higher payoff when the market is going up, so its covariance in that situation is higher, this particular type of asset does not need the same premium as the other one.
This particular condition it is called upside risk and the main measure for computing that us the upside beta (or β^+) that is calculate with the previous formula.

Another way to estimate the downside and upside betas which I consider in my practical application was the one used by Victoria Dobrynskaya in her research paper called: "Downside Market Risk of Carry Trades".
The asset class focus of this research is the currency market, by the way the logic could be extended to all market. The process of the estimation of upside and downside betas are made by a system of equations with the ordinary least square method (OLS). All are time series regressions made with the auxiliary help of an important dummy variable:
The dummy can take value zero or one based on the following conditions:

$$dummy_t = \begin{cases} 1, & if\ r_{MKT,t} > 0 \\ 0, & if\ r_{MKT,t} < 0 \end{cases}$$

+

So, the regression following the sequent formula:

$$r_{i,j} = \alpha_i + \beta_i r_{MKT,t} + \sigma_i dummy_t * r_{MKT,t}$$

Where the downer letter i, stands for the asset of the regression taken into consideration. Given that formula and the implication of the dummy variable the betas are the following:

- $\beta_i^+ = \sigma_i + \beta_i$, this represents the upside beta.
- $\beta_i^- = \beta_i$,

Starting with the Arrow-Pratt definition of risk aversion which is equal to $-\frac{u(w)''}{u(w)'}$,

Were u(w)'' stands for the second derivative of the utility of the wealth and u(w) of the first ones. Harvey and Siddique have shown that the investors do not only care about the mean and the variance of the returns when they are dealing with the construction of their portfolio and their concept is consistent under the Arrow-Pratt risk aversion definition.

This important financial notion implies a particularly corollary that is applied by risk-adverse investors, which is that there must have a non-increase absolute risk aversion so the second derivative of the utility of the wealth must be greater than zero.

The economists as previous mentioned demonstrated their hypothesis with the help of an important statistical measure which is the skewness. This conceptually compute the asymmetry of an analysed distribution; in the simpler formula it is:

$$skewness = E[(x - \mu)^3]$$

Where, μ stands for the mean of the distribution data, by construction if a distribution is perfectly symmetric in its values the skew must be equal to zero. For the asymmetric ones instead, it is less or greater than zero in respect to the direction of the so called "long tail" , for being positive it must be in the positive direction, negative in the opposite.

The average absolute deviation from the sample mean and other measures are used too even though less frequently than the standard deviation. Another interesting element regards the shape of the distribution of values: an example can be given by samples of income or expenditure data. These ones tend to be highly skewed, while financial data such as asset returns and exchange rate movements are relatively more symmetrically distributed but are also more widely dispersed than other variables that might be observed.

The skewness can be defined also with this other statistical formula:

$$skewness = \frac{\sum_{i=1}^{n}(x_i - \mu)^3}{s_x^3(n-1)}$$

The denominator n-1 represent the degree of freedom, s_x^3 is a measure of dispersion of the variable around the mean.
Back to Harvey and Siddique study specifically, when they are studying the unconditional distribution of their portfolio assets returns. So, with every traditional factor being equal, they are looking for portfolio which are more right skewed. In particular, assets that have a low skewness (so the ones that are more left skewness) are going to decrease the average portfolio skewness requires more higher returns.
The skewness can be defined starting by the classic asset pricing theory, where the basic first order condition for an investor in one period for holding a risk asset is equal to:

$$E\left[\left(1 + R_{a,t+1}\right)MRS_{t+1}\middle|\Omega_t\right] = 1$$

Where Ω_t represent the information set that has the investor in period t,
$\left(1 + R_{a,t+1}\right)$represent the total return on asset a. MRS_{t+1} is the marginal rate of substitution between t and t+1 of the investor, this could be seen as the stochastic discount factor that is available and discount all the risky asset payoffs.
With the no arbitrage condition this must be greater or equal than zero. The details of that must be declared in the assumptions about the distributions of the proxies and preferences of the investor. In the Capital Asset Pricing Model, it is imposed as linear in the market return.
It is important to highlight that most of the specifications of asset pricing fields can be seen as approximations that can improve the precision of the MRS (or stochastic discount factor).

In their study, the economists considered a quadratic form of the stochastic discount factor this because a linear form of the stochastic discount factor, as the Capital Asset Pricing Model, requires some important assumptions which are not applicable in a cross-sectional regression of the excess return on the betas.
The standard formula of CAPM is:

$E_t\left[r_{i,t+1}\right] = \beta_{i,t}E_t\left[r_{M,t+1}\right]$,

Where beta is equal to $\frac{Cov_t\left[r_{i,t+1}, r_{M,t+1}\right]}{Var_t\left[r_{M,t+1}\right]}$

Where r stands for the excess return of the asset I or the market with respect to the conditionally risk-free return.

There are some important assumptions that in the Capital Asset Pricing Model framework should be applied, in particular in the time series regression of the excess return over the risk-free rate of the asset on the relative market excess return there are some statistical results that should be respected. Most of all, the risk premium of the market should be equal for all the assets, there must be a zero intercept, all of the betas must be significant.
Instead, if we are dealing with a cross-section regression of the excess returns on the betas, as our case, the most important statement when we work with that statistical passages are that the betas, the market risk premium and all the slopes all together must be different from zero.
The economist instead chooses a quadratic form of the marginal rate of substitution in the market return framework, for many reasons, not only for the demonstration given before.
This also because you can deal with non-linear functions but mainly since this utility form respects the non-increasing absolute risk aversion. This important property states that if wealth goes up the risk aversion should not increase as well.

The quadratic form for the marginal rate of substitution implies an asset pricing model where the expected excess return on an asset is determined by its conditional covariance with both the market return and the square of the market return (conditional coskewness).
The conditional coskewness is defined as an asset covariance with the square of the market returns or in formula as:

$$coskew = \frac{E[(r_i - \mu_i)(r_{mkt} - \mu_{mkt})^2]}{\sqrt{var(r_i)}\ var(r_{mkt})}$$

Where r_i is the excess return over the risk-free rate of the asset I and μ represent the relative mean of the variable in consideration.
The basic quadratic form of the stochastic discount factor in formulas is these one:

$$m_{t+1} = a_t + b_t R_{M,t+1} + c_t R_{M,t+1}^2$$

Given the important assumption of the existence of a conditionally risk-free asset you could arrive to the important improvement of Kraus and Litzenberg that in 1976 made their version of the three moment Capital Asset Pricing Model.
This result in formula can be called a conditional version of their past mentioned model:

$$E_t\left[r_{i,t+1}\right] = \gamma_t E_t\left[r_{M,t+1}\right] + \delta_t E_t\left[r_{M,t+1}^2\right]$$

Where γ_t and δ_t are two parameters whose depends on four important market variables: skewness, coskwness, covariance and variance.
Harvey and Siddique shown that this model with the additional contributions of asset pricing factors discovered and documented at that moment could explain with a great success the asset pricing puzzle.

The final important relation before comes from the following mathematical passages:

$$E_t[r_{i,t+1}] = a_{1,t}Cov_t[r_{i,t+1}, r_{M,t+1}] + b_{2,t}Cov_t[r_{i,t+1}, r_{M,t+1}^2]$$

The central statement is that $a_{1,t}$ and $b_{2,t}$ are different from zero and the same among all assets in a cross-section of them. These two variables are functions of the four parameters highlighted before and are equal to:

$$a_{1,t} = \frac{Var_t[r_{M,t+1}^2]E_t[r_{M,t+1}] - Skew_t[r_{M,t+1}]E_t[r_{M,t+1}^2]}{Var_tE_t[r_{M,t+1}]Var_t[r_{M,t+1}^2] - (Skew_t[r_{M,t+1}])^2}$$

$$b_{2,t} = \frac{Var_t[r_{M,t+1}]E_t[r_{M,t+1}^2] - Skew_t[r_{M,t+1}]E_t[r_{M,t+1}]}{Var_tE_t[r_{M,t+1}]Var_t[r_{M,t+1}^2] - (Skew_t[r_{M,t+1}])^2}$$

Given the important assumptions and mathematical simplifications we could arrive to the important previous formula which proves the importance of the four market variables in particular of skewness and coskewness, whose are the main pillars of the economists' research and of my interest in this thesis work.

Back to the main principal idea, if we consider a two-period economy, we could do a very important Taylor expansion in order to explain the role of skewness and coskewness.
We start from the formula of the stochastic discount factor, in quadratic form with respect to the market return, from period t to t+1, and with the expansion we arrive to that result:

$$m_{t+1} = 1 + \frac{W_tU''(W_t)}{U'(W_t)}R_{M,t+1} + \frac{W_t^2U'''(W_t)}{2U'(W_t)}R_{M,t+1}^2 + o(W_t)$$

Where $\frac{W_t U''(W_t)}{U'(W_t)}$ is the relative risk aversion, and it is important to remind that the term must be negative, this is an indirect relation, explained in directions undelight that if the market returns will go up the MRS will go down. This concept is demonstrated by the decreasing marginal utility and, as written before, is consistent under the Arrow definition of a risk adverse investor, which is the main pillar of the market assumptions. $o(W_t)$ stands for the reminder of the expansion (and usually is statistically insignificant).

Kimball in 1990 paper matched the risk aversion with the prudence concept that all investors should have in their strategies, this is related to the precautionary savings manner that mostly of individual have . The arrow-pratt corollary of the non-increasing absolute risk aversion it involves that assets that increase the total portfolio skewness are always preferred by an investor. The directly consequence of that statement is that if an investor buy an asset with negative coskewness his portfolio will be more negatively skewed, so that asset must have a higher expected return with respect to the other assets.

As a consequence of this in a cross section of assets, it is assumed to be negative the slope of excess expected return on conditional coskewness with the market portfolio.

Thus, the premium for skewness risk over the risk-free asset's return assuming that the risk-free asset possesses zero betas with respect to all the factors being examined to explain the cross section of returns should also be negative.

In the previous demonstrations it is possible to decompose contributions of conditional covariance and coskeweness with the market to the expected excess return of a specific asset. Whereas this decomposition cannot be provided by alternative nonlinear frameworks such as Bansal and Viswanathan (1993).

Another possibility is added by skewness. It is the fact that the expected return needs to be higher to get investors hold low or negatively skewed portfolios since, at any level of variance, there is a negative trade-off of mean return and skewness.

The capital market "line" starts out at zero variance–zero skewness. Thinking of a ray from the risk-free rate (at zero variance) tangent to the surface at a particular variance- skewness combination, for that level of variance, there are many possible portfolios with different skewnesses. The point with the highest skewness is the tangency one. Then, adding another ray from the risk-free rate that is tangent to a different variance-skewness point there is a single efficient risky-asset portfolio, in the usual mean variance analysis. But in the mean-variance-skewness analysis there are multiple portfolios. The tangency of the investor's indifference surface to the capital market plane indicates the optimal portfolio for the investor.

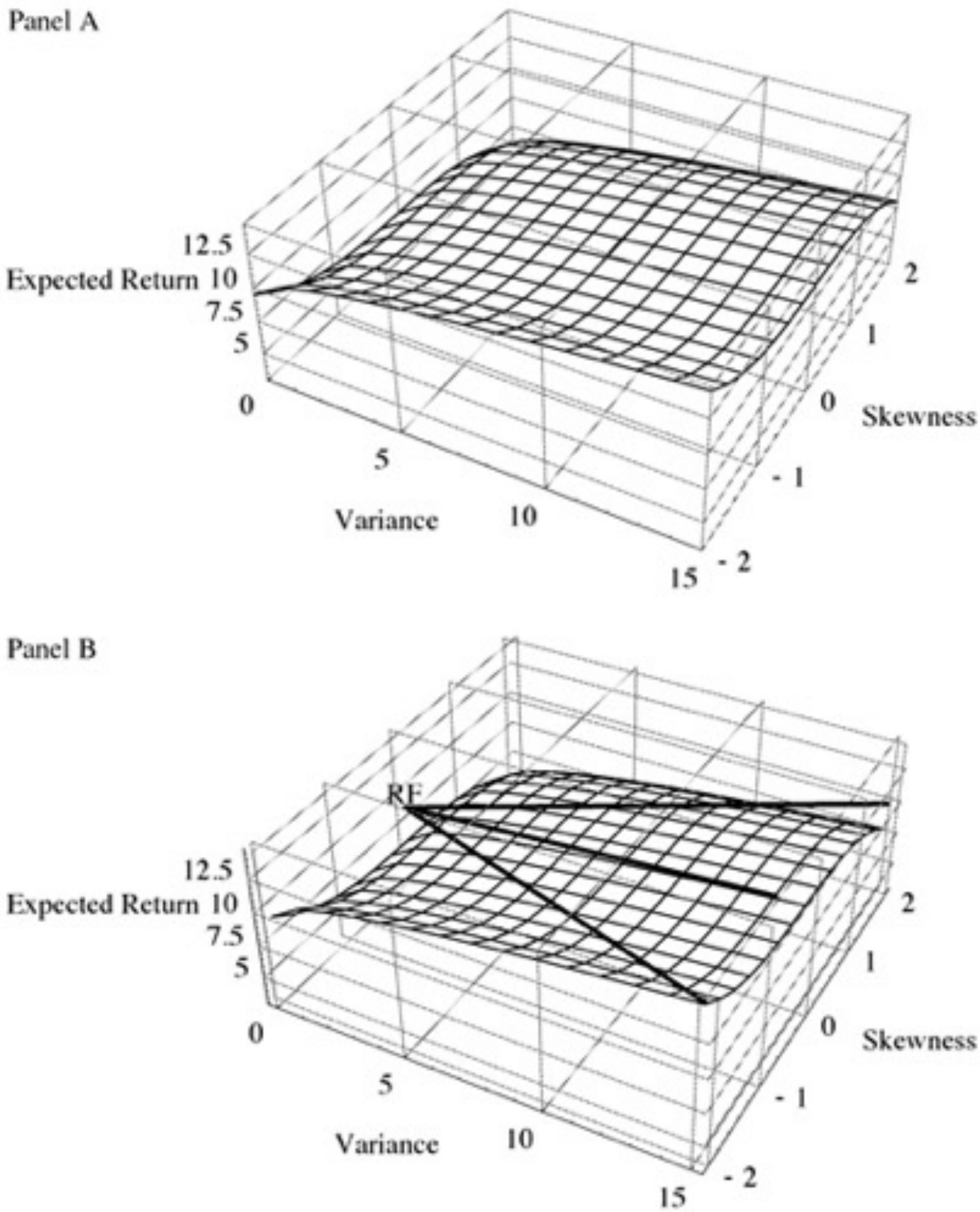


Figure 1: From Harvey and Siddinque. "Conditional Skewness in Asset Pricing Tests". The Journal of Finance (Vol LV, No. 3), June 2000, page 1271.

These pictures show a tridimensionality representation of the following measures: mean, variance and skewness.

The graphs represent the trade-offs between the variables. The data are taken form the Harvey and Siddique sample. The surfaces are generated using a positive trade-off between mean and variance and a negative trade-off between mean and skewness. Panel A presents the sur- face without a risk-free rate. In Panel B, rays are drawn from the risk-free rate to be tangential to the surface. The tangent points represent efficient portfolios.

Chapter 3 – Data and Methodology

For my research I took two datasets in order to doing my test, the difference between the two are the frequency of the data and the period took in consideration.

In particular for the first analysis, which has a focus on Multifactor and how the explain the funds return in the extended period, I decided -as it is a common practice in empirical research- to work with monthly returns, so I downloaded the monthly Fund Net Asset Value for my data from 29/02/2000 to 31/05/2022 from Bloomberg.

Secondly, I decided to analyse in more detail how it developed in Covid-19 period at my sample, so I restricted the analysis from 01/01/2020 to 31/03/2020 with a daily frequency.

I decided to work with that period because it was commonly used in other financial searches.

In order to decide my sample, I imposed some restriction in the Bloomberg Fund Screening which are the following:

- Take only the fund which are currently active.
- The fund primary class is share with an asset class allocation of equity, which I imposed must be minimum equal to 70% of the total fund value as I want to work with equity mutual funds.
- Take as geographic focus the nations which are considered by AQM as participants in the formations of their European Multifactor dataframe so they are: Belgium, Austria, Switzerland, Denmark, Germany, Finland, Spain, France, Greece, United Kingdom, Italy, Ireland, Israel, Norway, Netherlands, Sweden and Portugal.

The returns on factors of interest are taken for AQR Capital Management Data Set which follows the calculus of factors with the methods that I have reported in the literature review. I specifically worked with monthly, or daily in respect of the work that I done, factors returns. The portfolio returns are computed following the procedures studied in: Frazzini and Pedersen (2014), Fama and French (1992, 1993 and 1996), Asness and Frazzini (2013) and Asness, Frazzini and Pedersen (2013).

As reference of the risk-free rate as all the literature review suggested, I considered the one-month U.S treasury bill rate.

I selected from the AQR Capital Management Data Set a portfolio aggregate return that is representative of Europe market, it is computed especially by weighting each country's portfolio by the country's total lagged (t-1) market capitalization.

I took this portfolio as input for the multifactor variables of my regression for doing the regressions for my entire sample of European Equity Mutual Funds returns.

I decided to make all my empirical work in Python programming language as I believe that is the most friendly user code which can help programmers to make their own personal code with the help of mathematical and statistical packages which in my opinion make more easier the work of the user as they are not so rigid in their use and an user can write in the best way he will decide his code.

I used the following packages:

- Numpy: is a package that is used to with arrays and in all linear algebra field.
- Pandas: it stands for "Python Data Analysis Library" it is the most used package for data analyst, in particular it is very useful when you have to work with multi variables and fast implement difficult calculous. It is built with the connection with numpy.
- Openpyxl: it is used to have a connection between excel and python, in my case it was useful to load the data into python and create excel outputs in order to look at all my data analysis results.
- Statsmodels: it is a big statistical package, which has a lot of functionalities, I mainly used for doing ordinary least square regressions.
- Math and statistics: they are basically mathematical/statistical packages which allows you to do the basic computations.
- Matplotlib: it is the mainly graphical python package, I used to in order to make graphs.

The first thing that I implemented in my code is a clean of my initial data as Bloomberg download all the funds that are in activity in my period of interest but not from the start to the end, so I constructed a function who help me in that problem who is the following.

```
def clean_initial_data(xlsx_in, sheet_in, xlsx_out):
    """Read xlsx file, clean data droping columns and save file.
    Return deleted column numbers"""
    df = pd.read_excel(xlsx_in, sheet_name=sheet_in)
    columns_to_delete = [
        enne
        for enne, (columnName, columnData) in enumerate(df.iteritems())
        if enne != 0 and (columnData.iloc[0] == 'PX_LAST' or
                          str(columnData.iloc[1]) == 'nan')
    ]
    df.drop(df.columns[columns_to_delete], axis=1, inplace=True)
    df.to_excel(XLSX_FILE_OUT)
```

The result of this function in the case of my monthly European Equity mutual fund sample started with 2500 funds and after the elaboration I ended with 550 data that I used in my tests. Instead, the covid data sample started with 2932 data and ended with 2489 active funds in all of my interest period.

The main aim of the Multifactor part is to statistically look if the factors' investing strategies are used by the mutual fund managers and in particular what is the relevance of those and how they can explain the funds returns. In order to do that (as I will explain later in the results chapter) I mainly looked at the statistical and economical relevance of each factor in every regression. The formula that I used for doing my test is the following one, and so for testing my hypothesis I used a seven-factor model in an ordinary least square framework:

$$r_{x,t} - r_{f,t} = \propto_x + \beta_{1,x} MKT_t + \beta_{2,x} SMB_t + \beta_{3,x} HML\_FF_t + \beta_{4,x} HML\_D_t + \beta_{5,x} UMD_t + \beta_{6,x} BAB_t + \beta_{7,x} QMJ_t + \varepsilon_{x,t}$$

The subscript x stands for the x fund in consideration in the regression

The code that I wrote for each of my sample is this one, the difference between the two-period analysis is only the input variables, but they follow the same procedures:

```
for(columnName, data) in fund_data.iteritems():
    cnt+=1
    if cnt<=2:
        continue
    f_data = data.iloc[1:269]
    f_ret = f_data.pct_change()
    f_ret = f_ret.iloc[1:268]
    f_ret = np.asarray(f_ret)
    fund_return = np.asarray([i for i in f_ret])-rf
    df=pd.DataFrame({'MKT': mkt, 'SMB': smb, 'HML_FF': hml_ff, 'QMJ': qmj, 'HML_D':
hml_d,'BAB': bab,'UMD': umd,'fund_return': fund_return})
    # Eliminate inf
    df.replace([np.inf, -np.inf], np.nan, inplace=True)
    # Drop rows with NaN
    df.dropna(inplace=True)
    #Regression procedure
    x=df[['MKT','SMB','HML_FF','HML_D','BAB','QMJ','UMD']]
    y=df['fund_return']
    x=sm.add_constant(x)
    result=sm.OLS(y,x).fit()
    if result.rsquared >= 0.6:
        find = False
        for elem in result.tvalues:
            if elem > 2 or elem < -2:
                find = True
        if find:
            estimations.append(result.params)
            n+=1
```

For the purpose of analyse the downside risk of my sample I followed some of the statistical measures which I explained in the literature review. In particular, I started with the computation of the downside and upside beta with the Dobrynskaya method.

For that entire downside risk part, I made my tests on the bigger sample of European Equity Mutual Funds, so from 29/02/2000 to 31/05/2022.
With the following ordinary least square method for each equity mutual fund of my sample:

$$r_{i,j} = \alpha_i + \beta_i r_{MKT,t} + \sigma_i dummy_t * r_{MKT,t}$$

So practically I started with the construction of the dummy variable which follow the following statement:

$$dummy_t = \begin{cases} 1, & if\ r_{MKT,t} > 0 \\ 0, & if\ r_{MKT,t} < 0 \end{cases}$$

The code I wrote for computing it is the following:

```
c_array = np.asarray([1 if b > 0 else 0
                      for a in mkt_array
                      for b in a])
```

c_array represent the array which contains the value 0 or 1 depending on the value of the Fama French MKT factor computed by AQR dataset monthly.
Later, I constructed I code which is very similar to the previous one used for the Multifactor Analysis because it follows the same statistical procedure, so I used the same python packages. With the specification that the exogenous and endogenous variables are different so, they are that ones:

```
df = pd.DataFrame({'mkt': x1_array, 'dummy': c_array, 'fund_return': found_ret})
x = df[['mkt', 'dummy']]
y = df['fund_return']
```

Later as the theory said I used the coefficients of the market variable and dummy variable for defined the upside and downside betas so they become:

- $\beta_i^+ = \sigma_i + \beta_i$, this represents the upside beta.
- $\beta_i^- = \beta_i$.

Then I wrote I code to compute the upside beta and downside beta with Bawa and Lindenberg method. The formulas as I wrote before following the classic Capital Asset Pricing formula in order to compute the beta with the central difference that if I want to compute the downside beta I took the data that are referred to the moment that the mkt average return is below its relative mean and in the case of upside beta the opposite.
So the central line of my code for this purpose are this ones:

```
df2 = pd.DataFrame({'mkt': x1_array, 'fund_return': fund_excess_return})
Beta down df = df2[df2['mkt'] < mkt_mean]
beta_down = st.covariance(f_down, m_down) / st.variance(m_down)
```

I did the same lines for upside beta with the difference that in that case for having the dataframe of beta up I did the following: Beta up df = df2[df2['mkt'] > mkt_mean]

To conclude I computed the coskewness of my sample of funds with the following formula:

$$coskew = \frac{E[(r_i - \mu_i)(r_{mkt} - \mu_{mkt})^2]}{\sqrt{var(r_i)}\ var(r_{mkt})}$$

Where the subscript i stands for European equity mutual fund of consideration and mkt represent the market. All the variables in consideration are computed in excess of the risk-free rate. I took as benchmark the US monthly t-bill rate because it is universally taken in convention when you want to compute the risk free. I made this decision especially because also AQR dataset do also this strong assumption, and so I work with those data, and I want a reasonable comparison. For market data I considered the ones of AQR, so with the methods and source that I told before.

The more important lines that I wrote for this aim are:

```
mu_mkt.append(mean_mkt_c)
fund_c = df3['fund_return']
fund_c = np.asarray(fund_c)
f_c = []
for be in fund_c:
   f_c.append(be)
f_c = np.asarray(f_c)
mu_f=f_c.mean()
```

```
        mu_fund.append(mu_f)
        coskew=((f_c-mu_f)*((m_c-
mean_mkt_c)**2))/(math.sqrt(st.variance(f_c))*st.variance(m_c))
        coskew=np.asarray(coskew)
        coskew_m=coskew.mean()
        coskew_mean.append(coskew_m)
```

Chapter 4 – Results and Interpretations

As I said before, for the multifactor investment analysis I did a series of ordinary least square regressions, in whose I consider as dependent variable the excess return of each European equity mutual fund over the one-month T-bill rate and as independent variables the monthly factors return of AQR dataset.

For interpreting the results and attest if my hypothesis that European equity mutual funds follows factor investing strategies, I have followed the two methodologies used by Eduard Van Gelderen and Joop Huiji in their paper: "Academic Knowledge dissemination in the Mutual Fund Industry: Can Mutual Funds Successfully Adopt Factor Investing Strategies?". There two approaches change on the incidence that they express, as it relies on fact that if a found is involved in a particular factor investing, that it has an empirical demonstration, in a statistical or economical way. For the classification, I considered that a fund has an economically significant exposure to a particular factor investing strategy if the result of the beta relative to our strategy of analysis in greater than 0.25, despite of the citated work I considered that in absolute value. This is due to the fact that significant negative value shows that the fund has following an opposite strategy in respect to the factor of consideration.

In contrast, I assumed that the factor has a statistical significance, so the fund has a statistically significant exposure to a particular factor investing strategy if the t statistic relative to the beta of our factor of interest is in absolute value greater than 2.

Obviously these two results are not exclusive and the presence of one does not exclude the other one.

When I made the regressions, I also imposed two relevant conditions in order to work with only significant data whose are:

- R-squared of the regression in consideration must be greater or equal than 0.6. The reason for that choice is that I want to look at funds who are involved in factor investing so that the factors variables could explain the fund excess return.

- T statistic of the entire regression must be greater or equal than 2 in absolute value, the aim of that is that I want to have only statistical relevant data in my analysed sample.

As I said before, I made two class of regressions based on different data input.

First of all, I did a monthly regression of my sample of European equity mutual funds from 29/02/2000 to 31/05/2022 on the monthly AQR factors.

The two central restrictions that I imposed before take the sample of from 550 funds to 404.

Then I printed the results for each factor, both numerical and graphical summary statistics, made in two previous citated dimensions (statistical and economical):

- First, I did a statistical filter which choose only the funds whose have a relative t statistic whose is or greater than 2 or less than -2. This allows me to watch the results of the funds whose have a relevant statical impact in factor investing.
- Then I choose the funds whose have an intercept which is greater than 0.25, so I could say that the funds are really economical involved in these strategies.

The results for MKT factors are the following:

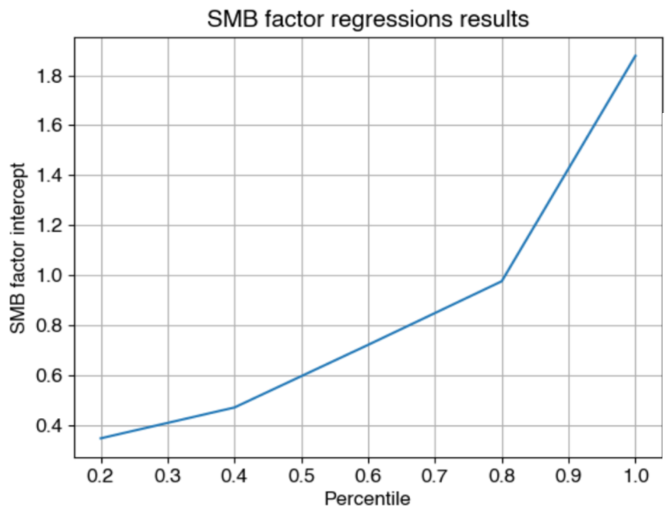

```
The statistical relevant MKT factor exposed funds number is:404
The economical relevant MKT factor exposed funds number is:400
The quintile of the MKT factor distribution is:
0.2    0.64960
0.4    0.70404
0.6    0.75542
0.8    0.82372
1.0    1.19200
```

Figure 2: MKT factor regression results   *Table 3: quintile of MKT factor distribution*

As I expected, almost every fund of my sample is exposed in both of the measures, as all of the funds invest in Europe equity markets so their result covary strongly with their environment results . For each factor I printed in the code also the percentile of the distributions of the fund's factors intercepts in order to have a summary of the economical exposure of my sample. I also plotted that in order to have a quick visual of the phenomenon.

The MKT results as I said are pretty obvious, in particular they are the bigger in terms of the number of funds that are exposed in this factor.

The second variable that I considered is the Small Minus Big factor which has a good respond in my big sample of data, as I saw that almost 70% of the sample is statistical involved in this strategy. This important result is given by the analysis of this factor, which represents the evidence that more than half of European equity mutual funds sample have invested in small capitalization stocks for obtaining the excess return over big stocks.




```
The statistical relevant SMB factor exposed funds number is:272
The economical relevant SMB factor exposed funds number is:226
The quintile of the SMB factor distribution is:
0.2    0.3485
0.4    0.4718
0.6    0.7222
0.8    0.9766
1.0    1.8781
```

Figure 3: SMB factor regression results *Table 4: quintile of SMB factor distribution*

The third variable that explain my mutual fund sample excess return over the risk-free rate is the High Minus Low of French Fama Factor. The results are quite satisfying as it can be stated that around 35% of fund managers decided to bid up on value stocks and follow investment based on a value-growth strategy. The economical relevance is slightly significant as the first quintile has an intercept of HML FF of 0.35.


```
The statistical relevant HML_FF factor exposed funds number is:143
The economical relevant HML_FF factor exposed funds number is:139
The quintile of the HML_FF factor distribution is:
0.2    0.34984
0.4    0.40734
0.6    0.46630
0.8    0.52584
1.0    0.77580
```

Figure 4: HML factor regression results *Table 5: quintile of HML factor distribution*

The other three factors give me unsignificant results as:

- High Minus Low Devil has only seventy intercepts are statistical relevant but only two funds deal with that strategy.
- Betting Against Beta shows me particular results as almost all of my sample intercept has a valuable t-statistic (for being precise 396) but only one has a quite significant intercept (0.4937)
- Up minus Down strategy has only 43 funds with a good t-statics and only three with a relevant intercept.

The quality minus Junk factor results underwrites the literature review results, as I could say that in the European equity mutual fund sample that I choose there is a good tendency on the fund manager on going long on quality assets and short on junk stocks. Around 35% of funds are statistical involved in this strategy, with ninety-four economical relevant results.

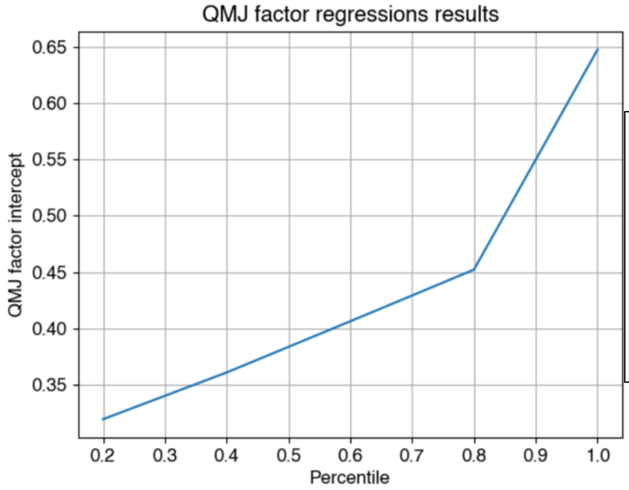


Figure 5: QMJ factor regression results

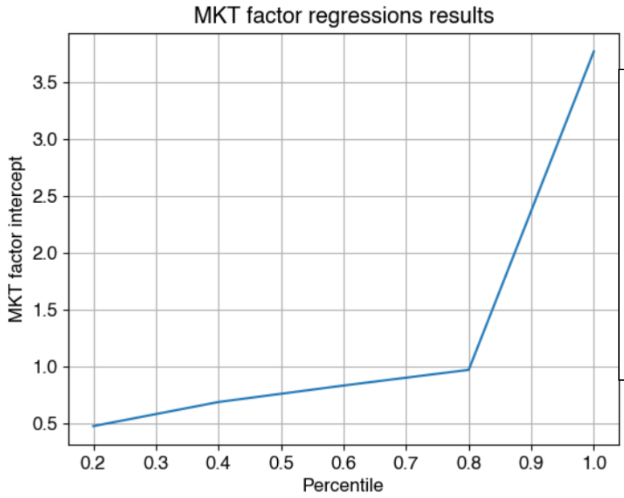
```
The statistical relevant QMJ factor exposed funds number is:141
The economical relevant QMJ factor exposed funds number is:94
The quintile of the QMJ factor distribution is:
0.2    0.31934
0.4    0.36084
0.6    0.40628
0.8    0.45210
1.0    0.64750
```

Table 6: quintile of QMJ factor distribution

I followed the same procedure for the Covid 19 sample of data, in particular I worked with the same criterion Bloomberg data research for having my sample of European equity mutual funds from 01/01/2020 to 31/03/2020 buy with a daily frequency. As I mentioned before, after the cleaning process which eliminates the funds whose are not active for all the period, I ended with 2489 active funds. Then, I did a resampling of that with my two important assumptions mentioned before on the results of each fund ordinary least square regression of their daily return in excess of the risk-free rate on the seven factor whose representing the factor investing strategy that I choose (so R-squared greater or equal than 0.6 and absolute value of t-statistic greater than 2). The first filter of my sample was that 1854 funds over the total number were involved in these strategies.

The aim of this test is to discover how factor investing worked in covid period and how important was for European equity mutual funds. So I did the same multi dimension analysis for all my seven independent variables in order to understand how much of the value of my dependent variable were explained by that.

Firstly, I analysed the importance of the Market Factor and as I expected it was statistical and economical important for almost of my data, as I could see from the summaries of the percentile of the distribution of all the MKT intercept of the regressions results and its graphical representations. So, I could say that all the funds are exposed on market movements, as it is explicit on their equity asset class focus.

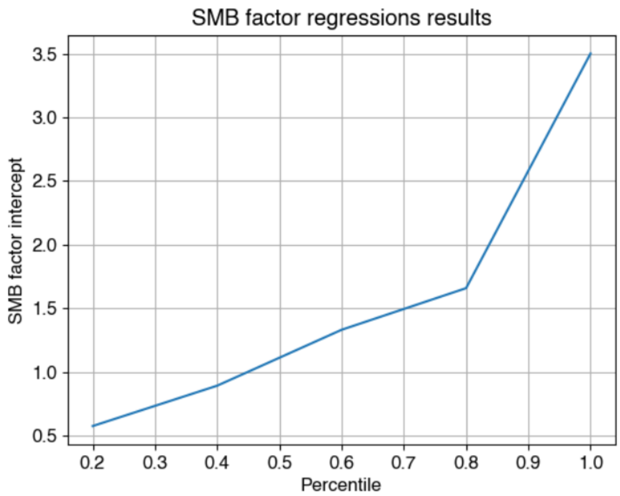



```
The statistical relevant MKT factor exposed funds number is:1847
The economical relevant MKT factor exposed funds number is:1732
The quintile of the MKT factor distribution is:
0.2    0.47402
0.4    0.68604
0.6    0.83120
0.8    0.96912
1.0    3.76940
```

Table 7: quintile of MKT factor Covid-19 distribution

Figure 6: MKT factor Covid-19 regression results

The Small Minus Big factor results represents an evident trend that was present in covid period of capture the size effect in their investment by the mutual fund managers. It is present big evidence that a lot of excess return of a consistent number of funds could be explained by the SMB factor intercepts as the average is near one. I could say that at least fifty percent of my sample of European equity mutual fund portfolios were involved in majority by small firms' stocks. However, the regressions show me that the t-statistic of the factor intercepts were near seventy percent of my restricted sample (after the assumptions).




```
The statistical relevant SMB factor exposed funds number is:1241
The economical relevant SMB factor exposed funds number is:933
The quintile of the SMB factor distribution is:
0.2    0.57436
0.4    0.89144
0.6    1.33064
0.8    1.65764
1.0    3.50140
```

Table 8: quintile of SMB factor Covid-19 distribution

Figure 7: SMB factor Covid-19 regression results

The third classic factor of analysis, so the French-Fama High Minus Low gave me not so convincing results, as about seven percent of my restricted grouping of funds were statistical and economical involved in this factor investing strategy. By the way I could say that the fund managers return who deal with strategy were very correlated with the classic value-growth factor because the mean of the HML founds intercepts was around two. Below I represented the results, the big size can be seen from the 80 percentile that is around 3.15.


```
The statistical relevant HML_FF factor exposed funds number is:120
The economical relevant HML_FF factor exposed funds number is:107
The quintile of the HML_FF factor distribution is:
0.2    1.54866
0.4    2.01954
0.6    2.44360
0.8    3.15414
1.0    5.58630
```

Table 9: quintile of HML French-Fama factor Covid-19 distribution

Figure 8: HML French-Fama factor Covid-19 regression results

The modified version of the previous written factor, so the High Minus Low Devil gave me not so great results as the statistical exposed funds were 72 and the economical only 30. Betting against beta factor and Quality Minus junk had strange results in my sample. As both type of strategies has statical relevant intercepts, 650 vs 503. But only few funds excess return over the risk-free rate has an economical relevant factor intercept, 10 vs 97 but with quite high values around 0.5 vs 0.8 mean.
The Up Minus Down factor regression gave me not so really important results are very low number of funds have significant intercepts.

Then I did the computations for the second argument of my thesis, so the Downside Risk phenomena. Firstly, I computed the downside and upside beta with the Professor Dobrynskaya method. So, I did an ordinary least square regression with a dummy variable for each mutual fund of my biggest period sample, so with monthly excess return over the monthly US T bill. In formula the following:

$$r_{i,j} = \alpha_i + \beta_i r_{MKT,t} + \sigma_i dummy_t * r_{MKT,t}$$

$$dummy_t = \begin{cases} 1, & if\ r_{MKT,t} > 0 \\ 0, & if\ r_{MKT,t} < 0 \end{cases}$$

- $\beta_i^+ = \sigma_i + \beta_i$, this represents the upside beta.
- $\beta_i^- = \beta_i$,

In this calculus I also applied my two strong assumptions that I used before, so I selected only the regressions which have an R-squared who is greater or equal than 0.6 and a total t-statistic whose is greater than two in absolute value. So, I end to work with a sample of 258 of European Equity Mutual funds.
In did this regression in order to empirically test if a high downside beta is correlated with high returns over time. This because the theory stated that the investors must be compensated for the risk embedded in the fact that the stocks who have a high downside beta present a stronger covariance with market if it is in a declining period. This aspect could be seen also in the construction of the dummy. So practically for each fund, I also computed the average of the excess return for all the data. Then I divided my sample in parts sorted in ascending order by the value of downside beta or upside beta (which I also denoted with Beta down and Beta up).
I divided the result into four partitions, which are not perfectly equal as my sample is not divisible for four, then for each important statistic I took the average and I represent the results in the following tables.

Sample partitions	Beta down	Beta up	Mean of excess returns
25.2%	0.67	0.49	0.15%
25.2%	0.48	0.32	0.21%
25.2%	0.89	0.68	0.22%
24.4%	1.04	0.76	0.39%

Table 10: Downside beta ascending order results computed with Dobrynskaya method

This table is formed by an ascending order for Beta down, the results verify the hypothesis that there must be a sort of downside risk premium, which is implicitly required by the investors in order to have in a portfolio stock who are characterized by a high downside beta.

Sample partitions	Beta down	Beta up	Mean of excess returns
25.2%	0.71	0.46	0.18%
25.2%	0.47	0.29	0.31%
25.2%	0.90	0.68	0.29%
24.4%	0.95	0.85	0.20%

Table 11: Upside beta ascending order results computed with Dobrynskaya method

This other table represent the summary results with the same methodology of the one before but with the distinction that this is formed with the criterion of an ascending order for upside beta. Since β^+ only measures exposure to a rising market, stocks that rise more when the market return increases should be more attractive and, on average earn low returns. I do not clearly observe this relationship in my sample as the funds whose I observe that have intermediary value of upside beta are the ones that have a higher mean of their relative excess returns. In other worlds, I do not observe a discount for stocks that have attractive upside exposure. So, my funds sample is inconsistent with the theory that the investors sentiment is usual very favourable toward assets that will have a high upside potential.

For having a comparison and deal more into in Downside and Upside betas computations I also computed the Bawa and Lindenberg version of the two measures. So, I took the same sample of data but I ended to work with more funds, because I did not take any assumption that has reduced my initial sample, so I empirically computed β^+ and β^- values of the excess monthly return over the monthly US T-bill of 550 European Equity mutual funds.
I removed two outliers so I ended with 548 funds.
As I mentioned before I follow the economists' formulas so:

$$\beta^- = \frac{cov(r_a, r_m | r_m < \mu_m)}{var(r_a, r_m | r_m < \mu_m)} \; and \; \beta^+ = \frac{cov(r_a, r_m | r_m > \mu_m)}{var(r_a, r_m | r_m > \mu_m)}$$

The monthly returns of my sample are 267 which after the computations is divided in 137 mkt data below its mean and 130 upside it.
As before I divided the two statistics in two sample sorted by ascending order of the relative value of the variable and matched with the funds excess returns mean. The aim of that is to

have a summary for testing the hypothesis behind the two statistics and measure their impact into the funds returns and strategy.

For the downside beta these are the results:

Sample partitions	Beta down	Mean of excess returns
25%	-8.03	-1.74%
25%	0.73	-2.46%
25%	0.85	-2.93%
25%	1.01	5.92%

Table 12: Downiside beta ascending order results computed with Bawa- Lindenberg method

The results do not reflect an increasing trend as I saw with the Dobrynskaya method, but I clearly say that in my results the highest values of β^{-} are always correlated with a very high excess return. So, it is partially verified that stocks with a high downside beta must carry a premium in order to entice fund managers to hold them in their portfolios.
Moreover for having another statistic that could confirm my idea, I computed the same calculus of the table but with fifty percent values of the β^{-}, the result are that the lowest part of the sample has a downside beta mean of -3.65 with a relative average excess fund returns of -2.10%, in opposite the upper part has an average β^{-} of 0.93 with the relative mean of the excess return of the equity mutual funds equal to 1.49%.
The table below reveals the same statistics but with respect to Upside Beta.

Sample partitions	Beta up	Mean of excess returns
25%	-0.24	2.75%
25%	0.53	2.76%
25%	0.65	3.16%
25%	0.80	3.59%

Table 13: Upside beta ascending order results computed with Bawa- Lindenberg method

The results quite shocked me as it seems that there is not present a discount for holding high upside beta in mutual funds' portfolios.

As I mentioned in the previous chapter I computed the Conditional Skewness of my sample with the following formula:

$$coskew = \frac{E[(r_i - \mu_i)(r_{mkt} - \mu_{mkt})^2]}{\sqrt{var(r_i)}\ var(r_{mkt})}$$

In particular I computed for each fund the monthly coskewness relative to the return over the risk-free rate. Then I did the average of the Coskew and the fund returns and I divided the results based on a Coskewness ascending order. Then I took all funds average results and I divided that in four equal parts based on ascending Conditional Skewness value.
So, I did the same work made for upside and downside betas and these are the results.

Partitions	Average Coskew	Mean of funds excess returns
25%	-0.58	0.38%
25%	-0.47	0.24%
25%	-0.40	0.17%
25%	-0.01	7.17%

Table 14: Coskewness ascending order results

These results have quite surprised me and I could say that the trend demonstrated in the literature review is not verified in my sample. The reason behind that may be that the market volatility in that period is high so the portfolios with low coskewness had low returns instead of the others or that these types of assets are basically characterized by negative coskewness over time so the theory trend is not verified.

Conclusion

The aim of this thesis was to fully analyze and describe the Multifactor in the European Equity Mutual Fund sector. For doing that after a in depth study of the literature review that treat the factor investing topics and downside risk phenomena, I build up some empirical test in order to analyze these topics into a representative sample of funds with the characteristics of my interest. This helps me to improve my data analytics skills especially in Python coding that I retain it could help me in my future working career. I started with the analysis of seven factors which they are recognized by the economists as the most important in the asset pricing theories, whose are the following: Market, Small Minus Big, High Minus Low, Up Minus Down, Quality Minus Junk, High Minus Low Devil and Betting against Beta. The hypothesis that I want to test with a series of ordinary least square regression was if that the mutual fund managers follow these theoretical strategies in their investment and if that who's of that is most followed and with what magnitude. I did this test for a sample of European Equity Mutual Funds monthly data of the Net Asset Value of each fund, made by Bloomberg data, from 29/02/2000 to 31/05/2022 which is a consistent time span to obtain results that are not characteristic of a single particular period but can be considered intrinsic and represented in my sector of interest. Then I computed the same calculus in the Covid-19 time period in order to test which strategies are followed in the investment of that particular era, so I considered the daily data for the same sample from 01/01/2020 to 31/03/2020.

Firstly, I did the twenty-year period ordinary least square regressions of each fund excess return over the risk-free rate on the seven factors of my interest. As I expected since the asset class of my interest is equity all of the sample is exposed on Market Factor. The Small Minus Big factor provides important findings in my sample as more than 70% funds are involved in this strategy with great betas results, I could say that in the investment strategies of these funds are betting on size effect when they are considering the assets who they want to invest into. High Minus Low results shown that around a third of the sample of funds decided to invest in firms that the stocks are characterized by having a low market price compared to their relative book value. The quality minus junk factor results underwrites that quality is searched in around 35% of my funds in their investments. Unfortunately, the other factors they did not gave me quite significant results.

The covid data gave me particular result, driven by the particular market period of interest. In opposite to the previous funding the Small Minus Big factor had a great respond in the

sample as the average betas were around one, this could be explained by the drastic market period that has awarded the small particular stocks with highest returns.
In the last part of my thesis, I treated the downside risk and I practically tested if the mutual funds of my sample cared about this risk aspect when they are doing their investment. I did some computations in order to demonstrated if the theories regarding this topic are verified in my data. In particular I want to show if a stock that has a high downside beta, so it is characterized by a strong covariance with the market if it is in a declining period it is rewarded by a premium during the other period in order to be maintained by the investor in his portfolio. For doing that I computed two important measures that are correlated with this phenomenon. The first one is the Downside Beta, and it is a particular beta that is formed when the stocks are in a period that market is showing returns which are below its mean.
I computed it with two different methods and both of them proved me quite good the theory finding that high downside beta is related with higher excess return over the risk-free rate of the respective stock, in my case fund return. The second measure is coskewness, or conditional skewness which is the asset covariance with the square of the market returns. The theory stated that usually a lower value of that must be associated with highest excess return over the risk-free rate of the asset in consideration. I computed for each fund the coskewness and the relative excess return but I did not find the theory statement. In my opinion this could be explained by the market volatility or by the type of my assets.

List of References

- Ang, Chen and Xing. Donwnside Risk, 2006. Published by Oxford University Press on behalf of The Society for Financial Studies.
- Asness, Clifford. S., Andrea Frazzini and Lasse H. Pedersen, 2013, Quality Minus Junk. Working Paper, New York University.
- Asness, Clifford, and Andrea Frazzini. "The devil in HML's details." The Journal of Portfolio Management 39.4 (2013): 49-68.
- Bawa and Lindenberg. "Capital market equilibrium in a mean lower partial moment framework". Journal of Financial Economics 5 (1977) 189-200.
- Black, Fischer, Michael C. Jensen, and Myron Scholes. "The capital asset pricing model: Some empirical tests." (1972): 79.
- Carhart, Mark M. "On persistence in mutual fund performance." The Journal of finance 52.1 (1997): 57-82.
- Dobrynskaya V. "Downside Risk of Carry Trades". Review of Finance 18(5), pp. 1885-1913, 2014.
- Fama, Eugene F., and Kenneth R. French. "Common risk factors in the returns on stocks and bonds." Journal of financial economics 33.1 (1993): 3-56.
- Fama, Eugene F., and Kenneth R. French. "Multifactor explanations of asset pricing anomalies." The journal of finance 51.1 (1996): 55-84.
- Frazzini, Andrea, and Lasse Heje Pedersen. "Betting against beta." Journal of Financial Economics 111.1 (2014): 1-25.
- Green W. "Econometric Analysis". 7th edition, Pearson Education Limited 2012.
- Harvey and Siddinque. "Conditional Skewness in Asset Pricing Tests". The Journal of Finance (Vol LV, No. 3), June 2000.
- Jegadeesh, Narasimhan, and Sheridan Titman. "Returns to buying winners and selling losers: Implications for stock market efficiency." The Journal of finance 48.1 (1993): 65-91.
- Nieto, Belén, and Gonzalo Rubio. "The Effects of the COVID-19 Crisis on Risk Factors and Option-Implied Expected Market Risk Premia: An International Perspective." Journal of Risk and Financial Management 15.1 (2022): 13.
- Data on factors: https://www.aqr.com/.
- Funds data acquired from Bloomberg data set.
- Filippo Ligabue. "Multifactor Analysis and Mutual fund Application". ICEF Term Paper.

Summary

This thesis took the basis from my term paper called: "Multifactor Analysis and Mutual Fund Application". The topic of my research is structured in two dimension whose are partially dependent: the first one is the multifactor part and the second one is the downside risk part.

The first goal of my thesis is to verify empirically how much the factor investing is important in mutual fund field and in particular in the strategies of the fund managers of the European Equity Mutual fund sector. I decided to focus my study on seven factors, starting with the classic three French Fama model factors (Market, Small Minus Big, High Minus Low) whose are recognized by the literature as the pillar of factor investing strategies. The other four factors are Momentum Factor, Quality Minus Junk, High Minus Low Devil and Betting Against Beta. I choose these factors because the literature review and the empirical evidence on this topic highlight that:

- investing with strategies that are focused on these factors in average generate time-varying expected risk premium;
- there is an important correlation between the returns over time of this strategies of stock picking and real activity;
- the selected factors are the more representative of all current asset classes and they are present not only in stock markets but also in the others.

After summarising the theories regarding the asset pricing factors, in order to prove my hypothesis, I developed some Python code to numerically verify if a consistent sample representative of the European Equity Mutual Fund managers invest in strategies based on those factors. For the literature review I started from the classic French Fama factors and then I decided to analyse also the more recent factors of Frazzini and Pedersen whose are punctually documented in the AQR Capital Management database.

The hypothesis that I want to test with a series of ordinary least square regression was if that the European equity mutual fund managers follow these theoretical strategies in their investment and if that who among them is the most followed and with what magnitude. I did this test for a sample of European Equity Mutual Funds monthly data of the Net Asset Value of each fund, made by Bloomberg data, from 29/02/2000 to 31/05/2022 which is a consistent time span to obtain results that are not characteristic of a single particular period but can be considered intrinsic and represented in my sector of interest. Then I computed the same calculus in the Covid-19 time period in order to test which strategies are followed in the

investment of that particular era, so I considered the daily data for the same sample from 01/01/2020 to 31/03/2020.

Firstly, I did the twenty-year period ordinary least square regressions of each fund excess return over the risk-free rate on the seven factors of my interest.

When I made the regressions, I also imposed two relevant conditions in order to work with only significant data whose are:

- R-squared of the regression in consideration must be greater or equal than 0.6. The reason for that choice is that I want to look at funds who are involved in factor investing so that the factors variables could explain the fund excess return.
- T statistic of the entire regression must be greater or equal than 2 in absolute value, the aim of that is that I want to have only statistical relevant data in my analysed sample.

Then I for each independent variable (factor), I imposed the two following conditions:

- First, I did a statistical filter which choose only the funds whose have a relative t statistic whose is or greater than 2 or less than -2. This allows me to watch the results of the funds whose have a relevant statical impact in factor investing.
- Then I choose the funds whose have an intercept which is greater than 0.25, so I could say that the funds are really economical involved in these strategies.

As I expected, since the asset class of my interest is equity, all of the sample is exposed on Market Factor. The Small Minus Big factor provides important findings in my sample as more than 70% of the funds are involved in this strategy with great betas results. So, I could say that in the investment strategies of these funds are betting on size effect when they are considering the assets who they want to invest into. High Minus Low results have shown that around a third of the sample of funds have decided to invest in firms that the stocks are characterized by having a low market price compared to their relative book value. The Quality Minus Junk factor's results underwrites that quality is searched in around 35% of my funds in their investments. Unfortunately, the other factors did not gave me quite significant results.

The covid data instead gave me interesting results, driven by the particular market period of interest. In opposite to the previous funding, the Small Minus Big factor had a great respond in the sample as the average betas were around one. This could be explained by the drastic market period that has awarded the small particular stocks with highest returns.

The second aim of my thesis consist in analysing the downside risk phenomena. This result was obtained due to a deep study of the most relevant theories that were written regarding this specific topic. I did some computations in order to demonstrated if the theories regarding this topic are verified in my data. In particular, I want to show if a stock that has a high downside beta -so it is characterized by a strong covariance with the market if it is in a declining period- is rewarded by a premium during the other period in order to be maintained by the investor in his portfolio. For doing that, I computed two important measures that are correlated with this phenomenon. The first one is the Downside Beta, which is a particular beta that is formed when the stocks are in a period where market is showing returns which are below its mean.

I computed it with two different methods (Dobrynskaya, Bawa and Lindenberg) and both of them proved me quite good the theory, finding that high downside beta is related with higher excess return over the risk-free rate of the respective stock -in my case fund return.

The second measure is coskewness, or conditional skewness, which is the asset covariance with the square of the market returns. The theory stated that usually a lower value of conditional skewness must be associated with highest excess return over the risk-free rate of the asset in consideration. I computed for each fund the coskewness and the relative excess return, but I did not find the theory statement. In my opinion, this could be explained by the market volatility or by the type of my assets.