# UNIVERSITA' DEGLI STUDI DI PADOVA


**DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"**

**DIPARTIMENTO DI SCIENZE STATISTICHE**


**CORSO DI LAUREA MAGISTRALE IN ECONOMICS AND FINANCE**

**TESI DI LAUREA**

**PERFORMANCE MEASURES: ANALYSING AND TESTING CORRELATION, STABILITY AND OTHER FEATURES BY MEANS OF A STUDY OF MANAGED PORTFOLIOS**


**RELATORE:**

**CH.MO PROF. Massimiliano Caporin**

**LAUREANDO: Riccardo Bado**

**MATRICOLA N. 1106279**



**ANNO ACCADEMICO 2017 – 2018**

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.
Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

Firma dello studente

___________________

# Table of contents

# Introduction

Nowadays many investors choose to devote part of their savings to **mutual funds**. The mutual funds are convenient because they allow investors with a moderate amount of capital to participate in the fate of financial markets with a well-diversified portfolio (ICI, 2016); they are also easily accessible. For this reason, in a few years of activity these financial instruments reached significant dimensions regarding both the number and type of funds available on the market, the number of investors and the size of the managed capital (Pianca, 2003). Consider a group of mutual funds in which a certain amount of capital may be invested in: a comparison among these potential investments might present certain difficulties.

A fundamental step in the management of a portfolio of funds is the **evaluation of the performance** of the investments, meaning to figure out whether the return is suitable in relation to the risk suffered (Plastira, 2014). For this reason, the increase in fund activities has affected considerably the interest in studies aimed to provide an accurate evaluation of these performances.

In many cases, this analysis is carried out through **performance indices** able to detect certain characteristics of the fund. These performance measures have a central role for investors, “allowing them to ex post compare the rankings of investment portfolios and to evaluate the real added value of managers” (Caporin, Jannin, Lisi & Maillet 2014).

William Sharpe, Nobel Prize winner and one of the originators of the Capital Asset Pricing Model, developing the Sharpe ratio in the mid-'60s was one of the precursors in the field of performance measurement. From that moment on, the **scientific research** on this topic became increasingly significant (Gottardo & Murgia, 1996). In fact, even though the Sharpe ratio may be a good measure with normally distributed returns, it could lead to incorrect investment decisions in the case of returns characterized by excess kurtosis or asymmetry (Zakamouline & Koekebakker, 2009). Therefore, over the years, a series of **different performance measures have been proposed** as an alternative to the Sharpe ratio, some of which aimed to solve the issue of non-normality of the return distributions.

In order to evaluate the performance of a mutual fund or, more generally, of a risky asset, it is known that the average return is not an exhaustive indicator. It is in fact necessary to take into account not only the **return** of the investment, but also the **risk** that it entails.

While everyone agrees on what the return is and how to measure it, the **risk** issue brings with it some complications. The idea of risk is less clear and presents ambiguous aspects.

Undoubtedly, it refers to the possible losses or loss of profits that an investment may involve (Pianca, 2003). However, the decision to measure the risk of a financial instrument with a single indicator turns out to be most of the time reductive, since it does not allow you to perceive every aspect of the concept of risk. Indicators of performance dispersion, such as variance, standard deviation or semivariance, are some of the tools commonly used to measure the risk of a portfolio (Bacon, 2011); however, many other indicators may be used.
As a result, more than 100 ways to measure portfolio performance are possible. The choice of a performance measure rather than another depends, *inter alia*, on the investor's preferences. However, in order to choose and understand how to properly use the right index, it is necessary to know some of the features of these measures.

**The aim of this study** is to compare the most influential performance measures, through both theoretical analyses and empirical applications. In particular, in order to provide a complete and comprehensive description, some specific features and properties of these indicators will be analysed throughout the paper.

**Chapter one** focuses on the analysis of the database of financial instruments used for the first part of the research. In particular, we are going to study certain characteristics of the selected funds, such as the normality of the returns distributions.
**Chapter two** focuses on the classification and on the description of the analysed performance measures. All these indices are defined and evaluated from a theoretical point of view.
**Chapter three** focuses on the analysis of certain features of the presented performance measures through an empirical investigation. The **first part** of the study regards an analysis of the possible **mutual correlation** between the indices. During this process, a static analysis of the rank correlation is firstly performed. Then we use a rolling approach in order to test the correlation between rankings over time. In the **second part**, the study focuses on the calculation of the **stability** of all the performance measures. Finally, the uncorrelated and more stable indices are combined in order to create a composite indicator that maximizes the stability over time.
**Chapter four** focuses on comparing portfolios built up following the same decision-making process but changing the performance measure that determines the asset allocation. The aim is to create funds of funds portfolios by rolling evaluations of every performance measure, analysing, at a later stage, certain characteristics of these portfolios.

Therefore, the **questions** now are: Are there groups of correlated indices? Which are the most stable performance measures? Does higher stability imply higher returns? Is there any measure which brings to a better fund selection?

# CHAPTER ONE

# 1 Analysis of the financial instruments selected for the study

## 1.1 Theoretical aspects of mutual funds

In order to analyse certain features of a set of performance measures we select a group of mutual funds with determined characteristics. A brief introduction on the **theoretical aspects of the mutual funds** is firstly presented.

### 1.1.1 Definition of mutual funds

The **mutual funds**, in the form of investment companies, are financial intermediaries that invest the money collected from a group of investors in a variety of financial assets, trying to create value (Lehman & Phelps, 2008). The funds are divided into single units with the same rights; each shareholder, therefore, participates proportionally in the gains or losses of the funds depending on the amount of shares owned (Pozen, Hamacher & Phillips, 2015). In the open funds, these units can typically be purchased or redeemed as needed at the price of the current fund's **net asset value (NAV)** (Sekhar, 2017); it is determined dividing the total value of the securities in the portfolio by the total amount of shares outstanding. Depending on the evolution of the prices of the underlying assets, the NAV may increase or decrease its value.
However, other factors influence the value of the NAV: the **ongoing charges**, for example, are fees that are directly charged in the NAV.

### 1.1.2 Mutual funds fees and expenses

In mutual funds, **the fees** can be classified into three categories (Borsa Italiana, 2017):

- ***Initial/redemption charges***: shareholders directly pay these fees when purchasing/selling the shares of the fund. Usually these fees are inversely proportional to the amount of the money invested (bigger the investment, lower the **percentage** charged). There are also funds whose shares may be sold without paying the commission, the so-called **no-load funds**.

- ***Ongoing charges***: they encompass the fund's professional fees, the management fees, the audit fees and the custody fees; they are charged as an annual percentage on the total assets under management. Reducing the value of the net assets of the fund, they directly hit the price of the NAV per share; therefore, all the investors of the fund pay, *pro rata*, these fees.
- ***Performance charges***: they may be charged in the event that the performance of the fund is exceeding a declared **threshold** (usually the **benchmark**). If that occurs, these fees are calculated on the amount of the investor's capital arising from this extra performance.

### 1.1.3 Classifications of mutual funds

Along with other relevant information about the fund, all these charges has to be clearly declared on the **KIID** (Key Investor Information Document), a mandatory information document for all the investment funds drawn up according to the rules of UCITS IV (Undertakings for Collective Investment in Transferable Securities) Directive (European Parliament and Council of the European Union, 2009). This document is subject to strict guidelines concerning the form, the content and the timing; it provides to investors information regarding investment objectives, risks, costs and historical performance (Rogo, 2013).
Depending on some specific feature, such as the asset allocation, the geographical exposure or the currency exposure, a fund can be classified in many different ways. A **classification of the funds** into uniform categories is crucial for the investors because it represents an important information for a first level evaluation of the financial instrument. In fact, it guides the investors among a large amount of assets, facilitating an initial screening of some characteristics of the fund, such as the asset allocation.
At the most basic level, mutual funds are organized into categories based on the **asset class** to which the underlying assets of the fund belong. In particular, they may be divided into the following groups: equity funds (those that invest predominantly in stocks), balanced funds (those that invest in both stocks and bonds), bond funds (those that invest predominantly in bonds), liquidity funds (those that invest predominantly in liquidity) and flexible funds. **Assogestioni** (Assogestioni, 2003) fixed some specific limit for all these classes, depending on the percentage of the fund portfolio invested in **equity**. In particular, liquidity funds cannot invest in equity, as well as the bond funds (with the exception of mixed bond funds that can invest maximum 20% of the portfolio in equity). Balanced funds can invest from 10% to 90%

of the portfolio in stocks, while equity funds has to invest at least 70% of the portfolio in equity. Finally, flexible funds do not have any kind of constraints in asset allocation.
Within each of these classifications, the funds may be classified into **further subcategories**, for example large/small capitalization fund for an equity fund, short/long term fund for a bond fund or, more generally, sector/thematic fund.

## 1.2 Database description

In this section, we describe the characteristics of the mutual funds selected **for the analysis**.

### 1.2.1 First funds selection

In this research, among the entire group of open fund distributed in Italy, we select the **pure balanced funds** (those that can invest from 30% to 70% of the fund portfolio in equity). The funds belonging to this category combine the high volatility of the equity market with the theoretically more stable returns of the fixed income market. We choose funds of the same category because, in this way, the performances of the analysed instruments are not totally driven by macro-movements of the equity/bond market but by the ability of the fund manager to create an extra return (compared to the funds of the same category) picking the single stocks or bonds. As a result, the performance measures should detect the differences in the funds management and not simply the movements of the market in which the funds invest.
In order to obtain from the analysis results as robust as possible, we use only funds with at least **10 year of history**.
In order to ensure that the effect of the exchange rate between different currencies does not influence the performances of the funds, we select all **Euro denominated** instruments. Bearing in mind that the exchange rate may vary significantly, the dynamics of its fluctuations cannot be underestimated.
From this first selection, we extract **45 funds**.

### 1.2.2 Analysis of Assets Under Management

We now present an analysis of the **total assets under management** of the 45 picked funds.

#### 1.2.2.1 AUM: definition and potential problems

The **AUM** is a measure of size of an investment firm (Haslem, 2010); it represents the total market value of the assets that the fund manages on behalf of investors. As examined before, funds charge their investors fees as a percentage of the total assets under management; thus, the AUM is one of the key factor for the revenue of the investment firms.
A **recent study** confirmed this assumption (Ibert, Kaniel, Van Nieuwerburgh, Vestman 2017). Analysing the revenues of around 500 Swedish **fund managers**, the authors observed that the principal factor to explain the differences in salary among them was not the fund performance, as someone could have thought, but the size of the AUM. On average, when a fund is significantly increasing its size, the increment of the fund manager revenue is 10 times bigger than the increment that he can obtain from higher performances. The performances, in fact, are not always able to influence the size of the fund. As stated also by this study, high performances of a fund in a year does not significantly affect the size of the fund in the **following** year. Investors, therefore, provide incentives to asset managers to act in their interests (Rajan, 2006), in order to increase the assets under management. However, because of the relevance of the AUM size for a fund manager, once that a fund collects a **great amount of assets** he may become more **risk averse**. If he is evaluated against his peers or a benchmarks, he may be induced to follow them (the phenomenon of “false funds active”, or “closet indexing”), taking similar positions while claiming to manage the fund actively (International Monetary Fund, 2015). Firstly, this action penalize the investors because they pay an active management receiving an expensive index tracker (Domian et al., 2015). Moreover, this behaviour can induce the transmission of shocks across assets (Broner, Gelos & Reinhart, 2006). See also Chakravorti & Lall (2003).
In addition to these problems, **a large amount of assets under management** may make a fund **difficult** and cumbersome **to manage**, hindering in this way the performance creation (Indro, Jiang, Hu & Lee, 1999). In fact, for example, investing a great amount of money in a single share can significantly affect its price. On the contrary, a limited amount of AUM allows the fund manager to move quickly in and out of stocks.
**Smaller funds** may also be **problematic**. First, they probably may have some problem to obtain a good diversification of the investments. Moreover, the fund expenses would tend to have a great impact on performance because of the difficulty to take advantages of economies of scale (Collins & Mack, 1997).

### 1.2.2.2 Comparison of AUM for the selected funds

For the reasons abovementioned, we avoid funds with an extremely big or extremely small AUM size.

In order to evaluate the size of the assets under management of a fund, it can be wise to make a **relative comparison** with the AUM of the funds of the same category.

Based on this consideration, we calculated for each of the 45 funds the **average assets under management** on all the history of the instruments **(Table 1)**. The AUM information, downloaded by the Bloomberg database, derives from a variety of sources including fund companies and third parties official documents.

***Table 1. AUM calculation for the first selection of funds***

| N° | ISIN | Fund Denomination | Average of AUM | Percentile |
|---|---|---|---|---|
| 1 | LU0095343421 | OYSTER MULTI ASSET DIVERSIFIED EUR | € 71.890.109 | 18,2% |
| 2 | LU0331284793 | BGF GLOBAL ALLOCATION C2 CAP. | € 17.168.118.889 | 95,5% |
| 3 | LU0099841354 | JB MULTICOOPERATION JB STRATEGY BALANCED (EUR) B | € 185.195.345 | 38,6% |
| 4 | LU0089291651 | PARVEST DIVERSIFIED DYNAMIC CLASSIC/CAP. | € 198.894.612 | 40,9% |
| 5 | LU0089650211 | SYMPHONIA LUX SICAV COMBINED DIVIDENDS DIST. | € 43.247.272 | 4,5% |
| 6 | LU0115099839 | JPM IF GLOBAL BALANCED D ACC. | € 506.314.660 | 68,2% |
| 7 | LU0132151118 | BNP PARIBAS L1 DIVERSIFIED WORLD BALANCED CLASSIC/CAP. | € 333.959.286 | 59,1% |
| 8 | LU0080749848 | FIDELITY FUNDS FIDELITY PATRIMOINE A ACC. | € 102.510.667 | 31,8% |
| 9 | LU0095623541 | JPM GLOBAL MACRO OPPORTUNITIES C ACC. | € 562.181.520 | 70,5% |
| 10 | LU0090850842 | LEMANIK EUROPEAN SPECIAL SITUATIONS A CAP. | € 50.231.185 | 9,1% |
| 11 | FR0010135103 | CARMIGNAC PATRIMOINE A ACC. | € 17.395.325.153 | 97,7% |
| 12 | FR0010434019 | ECHIQUIER PATRIMOINE | € 624.055.311 | 75,0% |
| 13 | DE0008478116 | DJE KAPITAL FMM-FONDS | € 430.453.872 | 65,9% |
| 14 | LU0056886558 | FIDELITY FUNDS FIDELITY PORTFOLIO SELECTOR MODERATE GROWTH A DI | € 226.423.894 | 45,5% |
| 15 | LU0052588471 | FIDELITY FUNDS EURO BALANCED A DIST. | € 706.221.673 | 77,3% |
| 16 | LU0212926058 | BGF GLOBAL ALLOCATION (EUR HEDGED) C2 CAP. | € 13.691.842.098 | 93,2% |
| 17 | LU0255639139 | NORDEA 1 STABLE RETURN AP | € 3.293.897.703 | 90,9% |
| 18 | LU0267387503 | FIDELITY FUNDS GLOBAL MULTI ASSET TACTICAL MODERATE A DIST. | € 210.876.568 | 43,2% |
| 19 | LU0247991317 | JPM IF GLOBAL BALANCED A DIST. | € 579.859.497 | 72,7% |
| 20 | LU0251130554 | FIDELITY FUNDS FIDELITY PORTFOLIO SELECTOR MODERATE GROWTH A A( | € 172.551.420 | 36,4% |
| 21 | LU0261950553 | FIDELITY FUNDS EURO BALANCED A ACC. | € 712.941.789 | 79,5% |
| 22 | FR0010306142 | CARMIGNAC PATRIMOINE E ACC. | € 19.776.155.453 | 100,0% |
| 23 | FR0010109165 | ODDO PROACTIF EUROPE CR-EUR | € 365.664.699 | 63,6% |
| 24 | LU0158187608 | AXA WORLD FUNDS C.TO GLOBAL FLEX 50 | € 91.605.007 | 27,3% |
| 25 | DE0009769893 | DWS VORSORGE AS (FLEX) | € 79.703.812 | 25,0% |
| 26 | DE000A0H0WT1 | AKTIVMIX VARIO SELECT | € 55.676.234 | 11,4% |
| 27 | DE0004156302 | SGR AKTIVMIX ERTRAG | € 36.961.194 | 2,3% |
| 28 | LU0346934713 | AZ FUND 1 C.TO ASSET POWER | € 336.240.966 | 61,4% |
| 29 | LU0134132231 | EUROFUND LUX C.TO IPAC BALANCED | € 23.278.554 | 0,0% |
| 30 | IT0003081525 | ALLIANZ MULTIPARTNER C.TO MULTI50 | € 66.376.744 | 15,9% |
| 31 | IT0000380060 | FONDERSEL | € 110.048.802 | 34,1% |
| 32 | IT0000380300 | EURIZON BILANCIATO EURO MULTIMANAGER | € 945.380.276 | 81,8% |
| 33 | IT0001080388 | EURIZON SOLUZIONE 40 | € 1.041.146.899 | 84,1% |
| 34 | IT0000380565 | EURIZON SOLUZIONE 60 | € 1.126.396.523 | 86,4% |
| 35 | IT0003677538 | UBI PRAMERICA GLOBAL MULTIFUND 50 | € 45.616.330 | 6,8% |
| 36 | FR0010376798 | FUNDQUEST SICAV C.TO BALANCED | € 74.612.306 | 20,5% |
| 37 | FR0010376822 | FUNDQUEST SICAV C.TO DYNAMIC | € 56.241.373 | 13,6% |
| 38 | FR0010607697 | GENERALI EQUILIBRE | € 259.598.902 | 50,0% |
| 39 | IE00B05MRN28 | MEDIOLANUM PORTFOLIO F. C.TO ACTIVE 80 | € 100.550.705 | 29,5% |
| 40 | IT0000380003 | ARCA BB | € 1.308.471.509 | 88,6% |
| 41 | IT0000382389 | FIDEURAM BILANCIATO | € 76.974.067 | 22,7% |
| 42 | IT0001051975 | FONDO ALTO BILANCIATO | € 256.685.509 | 47,7% |
| 43 | IT0003242366 | UBI PRAMERICA PORTAFOGLIO DINAMICO | € 267.492.498 | 52,3% |
| 44 | LU0121216955 | NN (L) PATRIM. SICAV C.TO BALANCED | € 301.303.249 | 56,8% |
| 45 | BE0159411405 | CANDRIAM SUSTAINABLE C.TO SUSTAINABLE MEDIUM | € 286.158.283 | 54,5% |

Source: own elaboration. Data obtained from Bloomberg Professional

Starting from these 45 funds, we thus **exclude** those in the upper 15% of the distribution and those in the lower 15% of the distribution.

### 1.2.3 Examined funds: the final selection

We then select only one fund for each investment company.

The **final 15 funds** are displayed in the following table **(Table 2)**, including a benchmark constructed as a composite index (Bacon, 2011).

***Table 2. Final selection of funds and benchmark***

| N° | ISIN | Fund Denomination | Currency | Fund Inception Date |
|---|---|---|---|---|
| 1 | LU0095343421 | OYSTER MULTI ASSET DIVERSIFIED EUR | EUR | 05/03/1999 |
| 2 | LU0099841354 | JB MULTICOOPERATION JB STRATEGY BALANCED (EUR) B | EUR | 30/07/1999 |
| 3 | LU0089291651 | PARVEST DIVERSIFIED DYNAMIC CLASSIC/CAP. | EUR | 30/01/1998 |
| 4 | LU0132151118 | BNP PARIBAS L1 DIVERSIFIED WORLD BALANCED CLASSIC/CAP. | EUR | 05/11/2001 |
| 5 | LU0080749848 | FIDELITY FUNDS FIDELITY PATRIMOINE A ACC. | EUR | 31/12/1997 |
| 6 | LU0095623541 | JPM GLOBAL MACRO OPPORTUNITIES C ACC. | EUR | 26/02/1999 |
| 7 | DE0008478116 | DJE KAPITAL FMM-FONDS | EUR | 17/08/1987 |
| 8 | DE0009769893 | DWS VORSORGE AS (FLEX) | EUR | 06/11/1998 |
| 9 | LU0121216955 | NN (L) PATRIM. SICAV C.TO BALANCED | EUR | 27/04/2001 |
| 10 | IT0000380060 | FONDERSEL | EUR | 27/08/1984 |
| 11 | IT0000380300 | EURIZON BILANCIATO EURO MULTIMANAGER | EUR | 25/03/1985 |
| 12 | BE0159411405 | CANDRIAM SUSTAINABLE C.TO SUSTAINABLE MEDIUM | EUR | 01/04/1996 |
| 13 | IT0000382389 | FIDEURAM BILANCIATO | EUR | 22/06/1987 |
| 14 | IT0001051975 | GIE ALTO BILANCIATO | EUR | 01/04/1996 |
| 15 | IT0003242366 | UBI PRAMERICA PORTAFOGLIO DINAMICO | EUR | 12/04/2002 |
| 16 | BENCHMARK | 50% BARCLAYS EUROAGG TR INDEX - 50% FTSE ALL WORLD INDEX | | |

Source: own elaboration

### 1.2.4 Benchmark and risk free rate

#### 1.2.4.1 Benchmark

Given that the analysed funds belong to the pure balanced category, the **benchmark** is 50% composed of **Bloomberg Barclays Euro Aggregate Bond Index** and 50% of the **FTSE All World Index**.

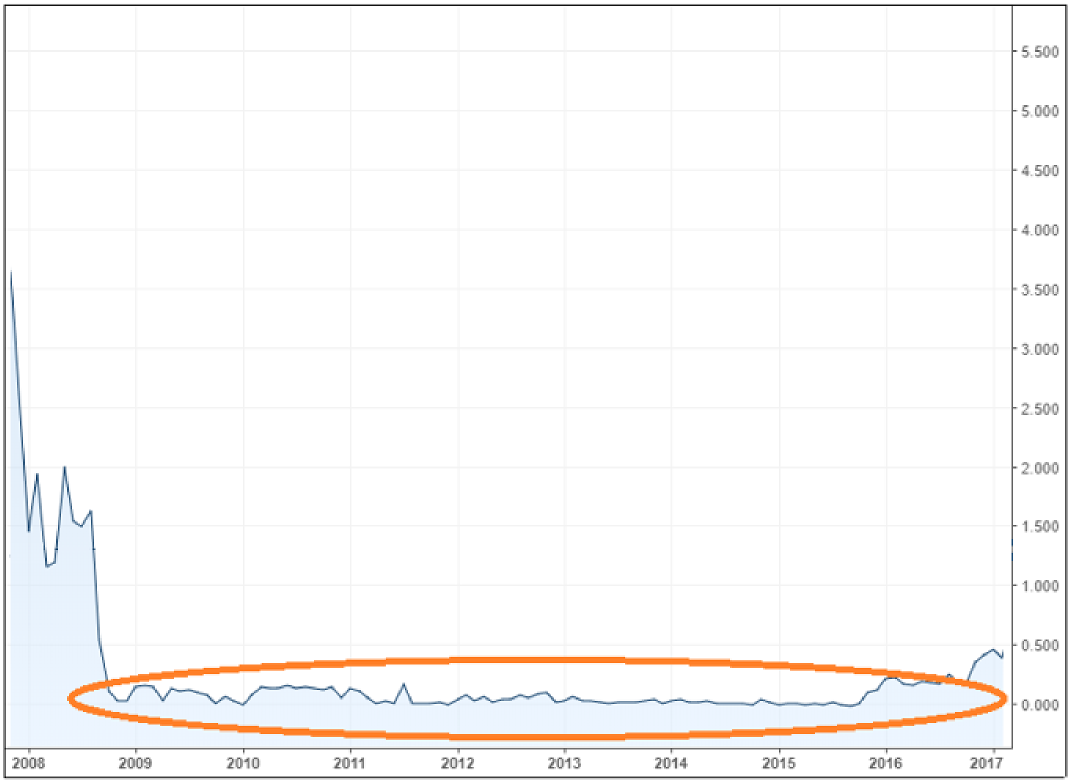
The Bloomberg Barclays Euro Aggregate Bond Index is an index that aggregates the euro denominated fixed-rate bond market, including treasuries, government-related, corporate and securitized issues.

The FTSE All World Index is a free float market capitalization weighted index representing the performance of the large and mid-cap stocks from the FTSE global Equity Index Series, covering both developed and emerging markets.

#### 1.2.4.2 Risk free rate

The risk-free rates are typically represented by the Treasury bills, which are assumed to have zero default risk because backed by the U.S. government (Bacon, 2011). As it can be seen from the following graph **(Figure 1),** representing the one-month US Treasury yield curve, in the last years, because of the quantitative easing of the Federal Reserve, the rate is often close to zero. For this reason, the **risk-free rate** is set to zero.

***Figure 1. Historical United States 1-Month bond yield curve***




Source: published on Investing.com, powered by TradingView

### 1.2.5 Conclusions

Summarizing, we choose to analyse pure balanced funds distributed in Italy, euro denominated, with at least 10 years of history, avoiding funds with extreme AUM and selecting maximum one fund for each investment company. Fifteen funds come out from this selection.

## 1.3 Real data analysis of the selected funds

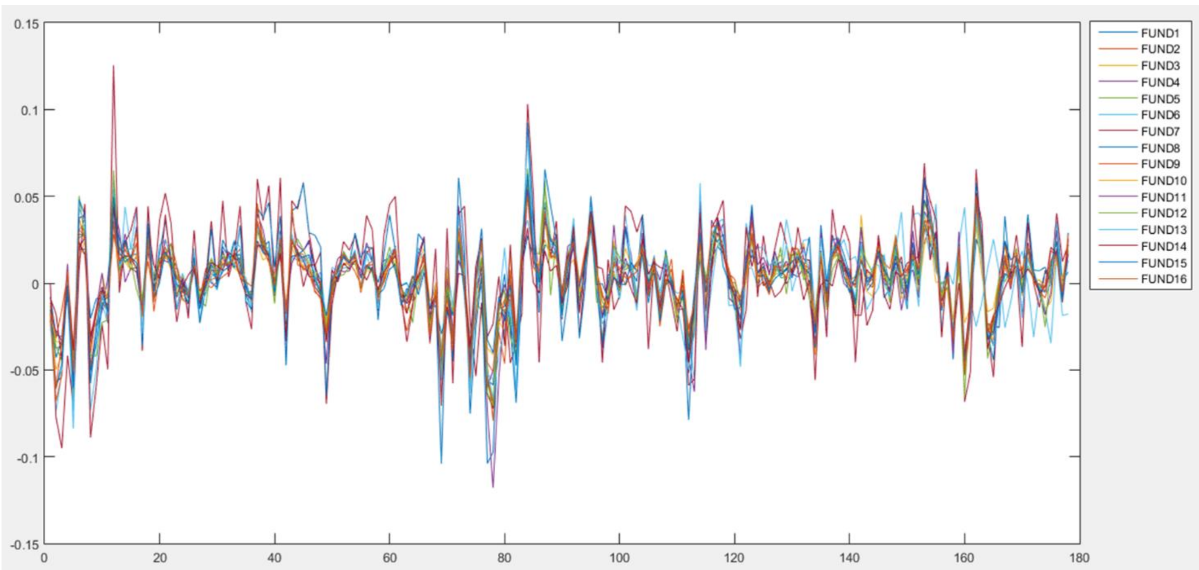
### 1.3.1 Time series and statistics about funds

The first date of common history of the funds is the **18/04/2002**; the final date of the analysis is the **28/02/2017**. For each fund, we thus obtain 179 monthly prices, equivalent to **178 monthly returns**.

It is also constructed a series of 60 months window rolling returns.

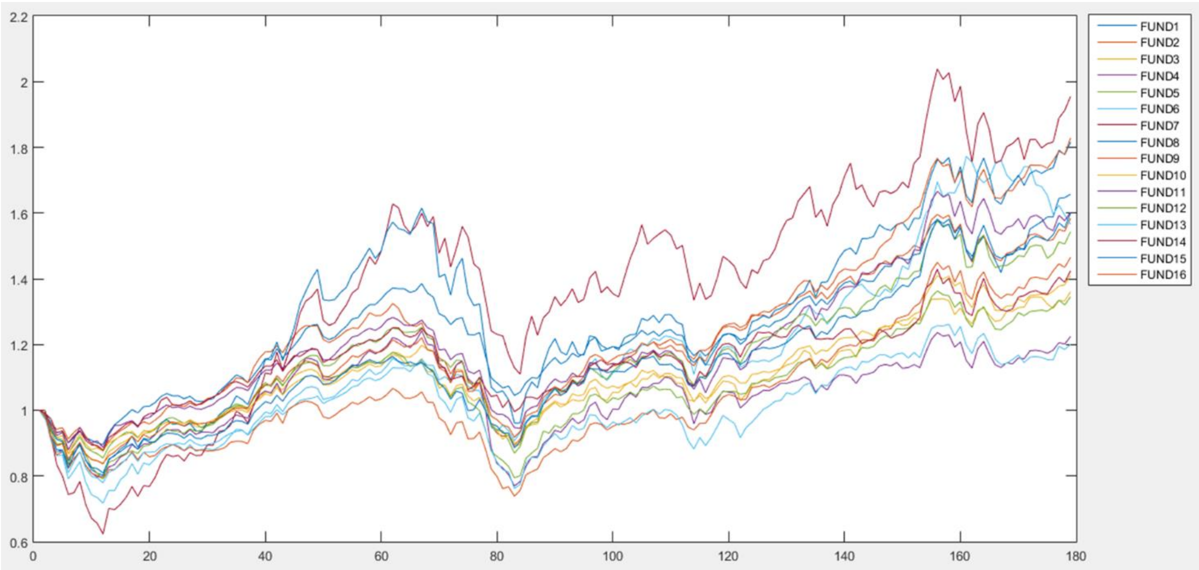
Now we present the **returns series** and the relative **price series** for the selected 15 funds and for the benchmark (FUND16 in the legend). **Figure 2** and **Figure 3** show the series for the monthly data, whereas **Figure 4** and **Figure 5** are referred to the 60 months rolling data.

***Figure 2. Time series of monthly returns***



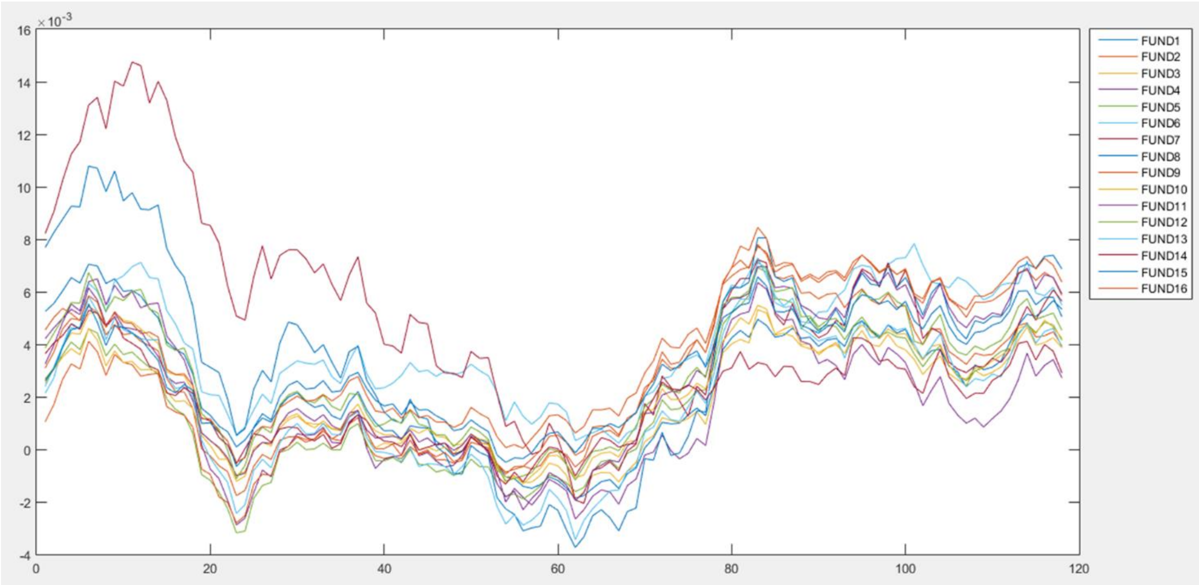

Source: own elaboration. Data obtained from Bloomberg Professional

***Figure 3. Time series of monthly prices***



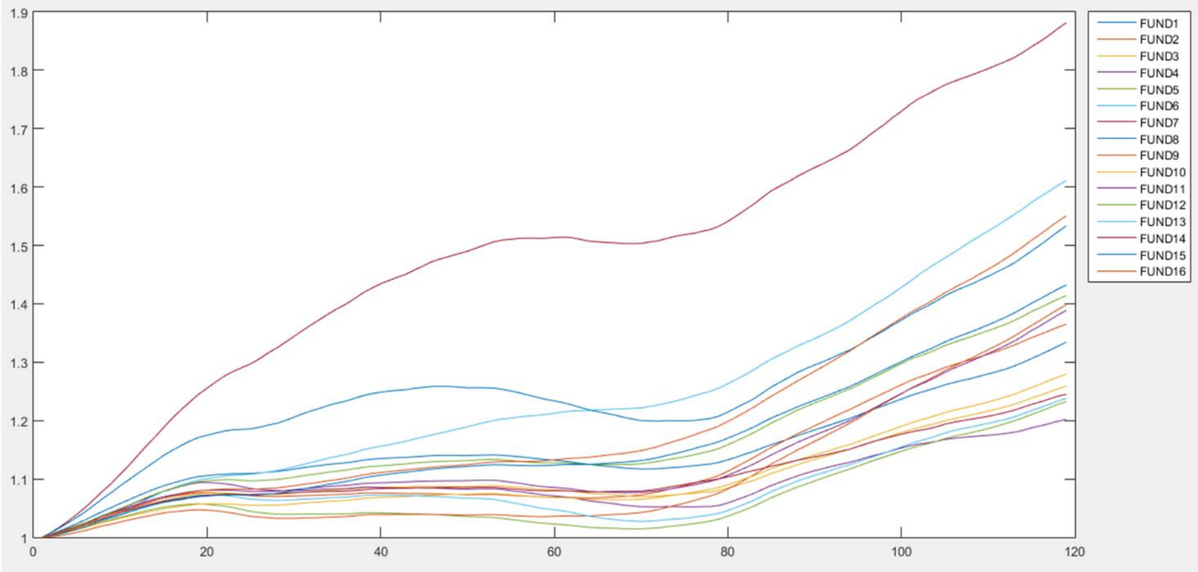

Source: own elaboration. Data obtained from Bloomberg Professional

***Figure 4. Time series of 60 months rolling returns***




Source: own elaboration. Data obtained from Bloomberg Professional

***Figure 5. Time series of 60 months rolling prices***




Source: own elaboration. Data obtained from Bloomberg Professional

Several results may be inferred analysing these figures. The greater stability of the rolling data rather than the monthly data is evident from a comparison of the price series. It is also clear that the funds belong to same category; in fact, investing in the same market, their paths end up getting similar in every figure.

In **Table 3** and **Table 4,** we present some **statistics** referred to, respectively, monthly data and rolling data. The volatility values are very different when calculated using monthly returns rather than rolling returns; the results confirm that the rolling data are more stable.

***Table 3. Statistics of funds calculated using monthly data***

| Fund | Mean | St. Dev. | Min | Max | Skew | Kurt | P-Value | JBStat |
|---|---|---|---|---|---|---|---|---|
| FUND1 | 0.300% | 1.755% | -5.861% | 4.141% | -0.778 | 0.957 | 0.15% | 24.752 |
| FUND2 | 0.277% | 1.918% | -7.922% | 5.342% | -0.849 | 1.831 | 0.10% | 46.217 |
| FUND3 | 0.192% | 1.936% | -6.642% | 5.612% | -0.711 | 1.088 | 0.17% | 23.799 |
| FUND4 | 0.147% | 2.567% | -11.785% | 5.845% | -1.004 | 2.277 | 0.10% | 68.326 |
| FUND5 | 0.186% | 1.970% | -6.990% | 6.604% | -0.388 | 1.340 | 0.37% | 17.784 |
| FUND6 | 0.282% | 2.448% | -8.383% | 5.760% | -0.712 | 1.104 | 0.16% | 24.092 |
| FUND7 | 0.445% | 3.697% | -9.507% | 12.544% | -0.260 | 0.150 | 27.94% | 2.177 |
| FUND8 | 0.385% | 3.117% | -10.401% | 9.226% | -0.795 | 1.580 | 0.10% | 37.269 |
| FUND9 | 0.240% | 2.246% | -6.913% | 6.189% | -0.731 | 1.034 | 0.17% | 23.791 |
| FUND10 | 0.209% | 1.770% | -5.755% | 5.201% | -0.596 | 0.779 | 0.58% | 15.018 |
| FUND11 | 0.288% | 2.115% | -5.969% | 5.742% | -0.409 | 0.636 | 2.49% | 7.964 |
| FUND12 | 0.270% | 2.293% | -7.105% | 6.496% | -0.657 | 1.437 | 0.10% | 28.123 |
| FUND13 | 0.134% | 2.493% | -7.359% | 8.406% | -0.500 | 0.778 | 1.02% | 11.916 |
| FUND14 | 0.219% | 1.987% | -6.543% | 6.908% | -0.204 | 0.888 | 3.17% | 7.091 |
| FUND15 | 0.286% | 2.103% | -6.081% | 6.292% | -0.385 | 1.097 | 0.78% | 13.333 |
| BENCHMARK | 0.359% | 1.964% | -6.069% | 5.038% | -0.761 | 1.063 | 0.14% | 25.536 |

Source: own elaboration. Data obtained from Bloomberg Professional

***Table 4. Statistics of funds calculated using 60 month rolling data***

| Fund | Mean | St. Dev. | Min | Max | Skew | Kurt | P-Value | JBStat |
|---|---|---|---|---|---|---|---|---|
| FUND1 | 0.245% | 0.241% | -0.186% | 0.705% | -0.034 | -1.135 | 3.85% | 6.358 |
| FUND2 | 0.264% | 0.262% | -0.176% | 0.780% | 0.076 | -1.415 | 1.61% | 9.964 |
| FUND3 | 0.196% | 0.209% | -0.195% | 0.532% | -0.059 | -1.410 | 1.65% | 9.846 |
| FUND4 | 0.156% | 0.258% | -0.288% | 0.650% | 0.255 | -1.120 | 2.87% | 7.450 |
| FUND5 | 0.178% | 0.255% | -0.318% | 0.692% | 0.011 | -1.323 | 2.15% | 8.606 |
| FUND6 | 0.405% | 0.214% | 0.024% | 0.784% | -0.042 | -1.312 | 2.20% | 8.503 |
| FUND7 | 0.538% | 0.382% | -0.206% | 1.475% | 0.553 | 0.105 | 4.20% | 6.059 |
| FUND8 | 0.364% | 0.384% | -0.374% | 1.078% | -0.219 | -0.932 | 5.51% | 5.208 |
| FUND9 | 0.285% | 0.304% | -0.281% | 0.845% | 0.176 | -1.390 | 1.56% | 10.107 |
| FUND10 | 0.209% | 0.192% | -0.131% | 0.549% | -0.069 | -1.364 | 1.88% | 9.235 |
| FUND11 | 0.279% | 0.281% | -0.230% | 0.696% | -0.075 | -1.471 | 1.37% | 10.750 |
| FUND12 | 0.294% | 0.239% | -0.120% | 0.711% | -0.102 | -1.367 | 1.82% | 9.395 |
| FUND13 | 0.182% | 0.268% | -0.344% | 0.714% | -0.138 | -1.200 | 2.86% | 7.456 |
| FUND14 | 0.186% | 0.169% | -0.133% | 0.530% | 0.044 | -1.132 | 3.87% | 6.338 |
| FUND15 | 0.305% | 0.199% | -0.065% | 0.657% | -0.106 | -1.307 | 2.15% | 8.622 |
| BENCHMARK | 0.373% | 0.248% | -0.049% | 0.772% | 0.062 | -1.525 | 1.18% | 11.504 |

Source: own elaboration. Data obtained from Bloomberg Professional

## 1.3.2 Study on the normality of fund returns distributions

Through a further analysis of these data, we can extrapolate other results, especially regarding the possibility to approximate the distributions of the return series to a **normal distribution**.

### 1.3.2.1 Literature on normal distribution as proxy for fund returns distributions

The model of Louis Bachelier (1900) for stochastic process became the prototype in modern finance for stock pricing processes. For a long time **Gaussian models** were applied in finance especially referred to stock, indices or funds returns. In fact, many stock valuation models has the normal distribution as main assumption: Markowitz Portfolio Theory (Markowitz 1952), Capital Asset Pricing Model (Sharpe 1964), Option Pricing Theory (Black and Scholes 1973). However, many financial economists, through many empirical studies, noticed that stocks returns, indices returns or funds returns are **badly fitted** by Gaussian distribution, mainly due to heavy tails and strong asymmetry (Ivanovski, Stojanovski, Narasanov, 2015). In fact, distributions of fund returns usually have negative skewness and positive excess kurtosis (fat tails) for several causes (diBartolomeo, 2014):

1) The “Central Paradox of Active Management”; 2) the distribution of security returns over short intervals; 3) the structurally short volatility of many funds.

1) ***The "Central Paradox of Active Management"***: all the active fund managers must believe their future returns will be above benchmark in order to pursue active management, but it is axiomatically true that roughly half of active managers produce below average results (diBartolomeo, 2010). This means that the investors have to suffer an additional risk referred to the high volatility of the returns around the mean (this additional portfolio risk is called "strategy risk" (Qian & Hua, 2005)).

2) ***Distribution of security returns over short intervals***: the frequency of large magnitude events in financial markets seems much greater than is predicted by the normal distribution (Mandelbrot, 1963), especially over shorter time horizons (a fund can enter and exit into a position very quickly).

3) ***Structurally short volatility of many funds***: many fund strategies are based on "value" or "momentum" driven security selection strategies. Momentum strategies, for example, buy on price strength and sell on weakness; this approach can exacerbate the large movements, causing the fat tails.

#### 1.3.2.2 Analysis of normality for the selected funds

In this section, we test the **accuracy** of **Gaussian distribution assumption** for our funds, an essential hypothesis for some performance measures. Normal distributions are symmetric and with a kurtosis equals to zero. When returns distributions take this form, the characteristics of a financial instrument can be measured with only two variables, the expected return and the standard deviation (Damodaran 2006).

**Kurtosis** (Kenney & Keeping, 1951) characterizes the relative peakedness or flatness of a distribution compared to the one of the normal distribution. This statistical measure has a significant importance for the investors, representing the possibility that prices change significantly.

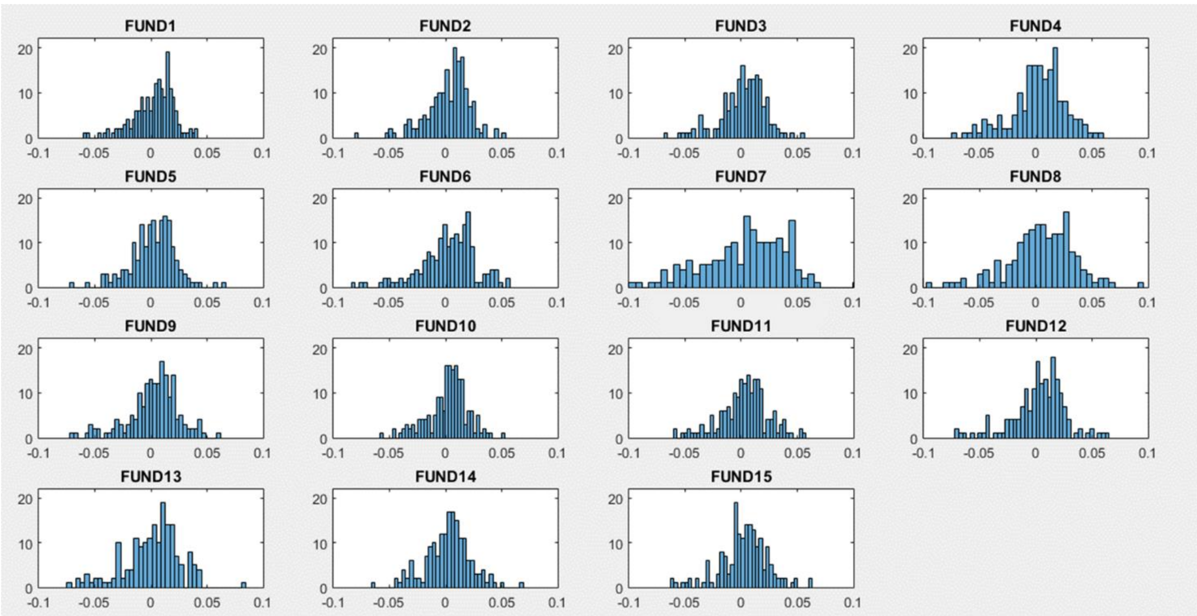
For a random variable x, kurtosis is defined as $Kurt(x) = \frac{E[(x-\bar{x})^4]}{\sigma^4} - 3$, where $E[(x-\bar{x})^4]$ is the fourth moment around the mean and $\sigma$ the standard deviation of x. Distributions with zero kurtosis are called mesokurtic, distributions with high kurtosis (heavy tails) are called leptokurtic, while distributions with negative kurtosis are called platykurtic (thinner tails and a flat top near the mean).

**Skewness** is a measure of symmetry, or, more precisely, lack of symmetry. It is defined as $Skew(x) = \frac{E[(x-\bar{x})^3]}{\sigma^3}$, where $E[(x-\bar{x})^3]$ is the third moment around the mean and $\sigma$ the standard deviation of x. If the skewness is positive (right skewed distribution), most values are concentrated on the left of the mean, with extreme values on the right. On the other hand, if the skewness is negative (left skewed distribution) most values are concentrated on the right of the mean, with extreme values on the left. Finally, if the skewness is equal to zero, the mean, the median and the mode corresponds to the "centre" of a set of data (Dean & Illowsky, 2017).

Regarding the monthly data, we can observe from **Table 3** that all the funds returns distributions have **positive kurtosis** and **negative skewness**. As a result, large changes in prices are much more common in funds returns than the normal distribution expects and the extreme values are concentrated on the left part of the distribution.

The histograms **(Figure 6)** confirms that all the returns distributions are leptokurtic and left skewed, with most values concentrated on the right of the mean and with extreme values to the left, with no exception.

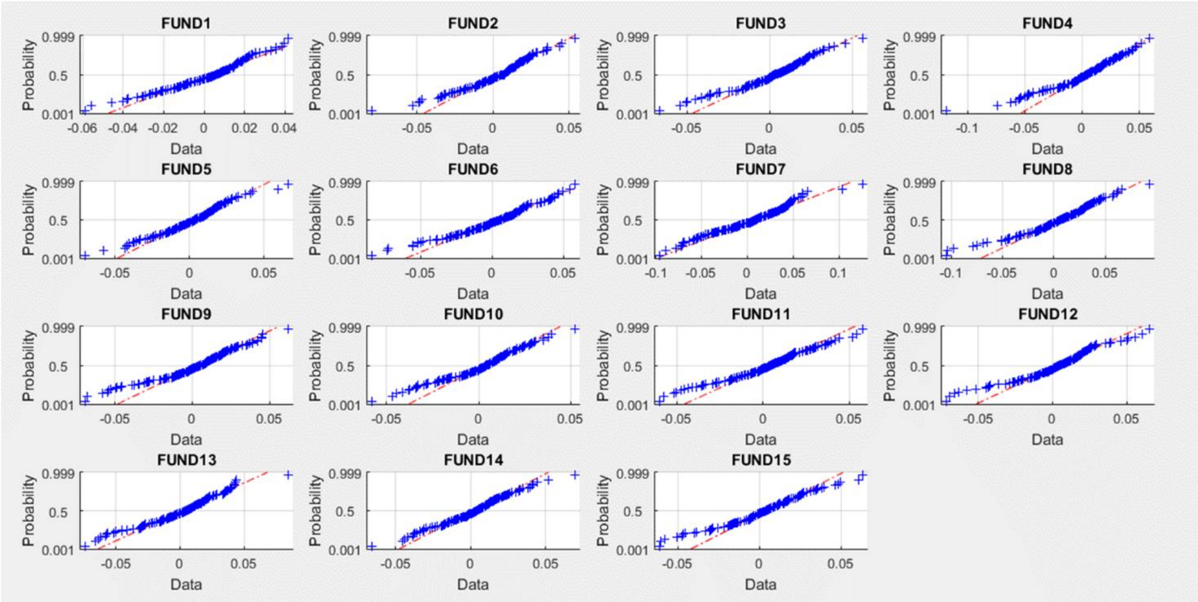
***Figure 6. Histograms of monthly returns***




Source: own elaboration. Data obtained from Bloomberg Professional

Despite the highlighted positive kurtosis and negative skewness, we present the "**Normal Probability Plots**" (Q-Q plot) of each fund in order to confirm that the returns distributions cannot be approximated by a Normal distribution. The **Figure 7** shows that almost all the distributions seem to deviate from the normal distribution, except for one case (FUND 7).

***Figure 7. Q-Q plots of monthly returns***

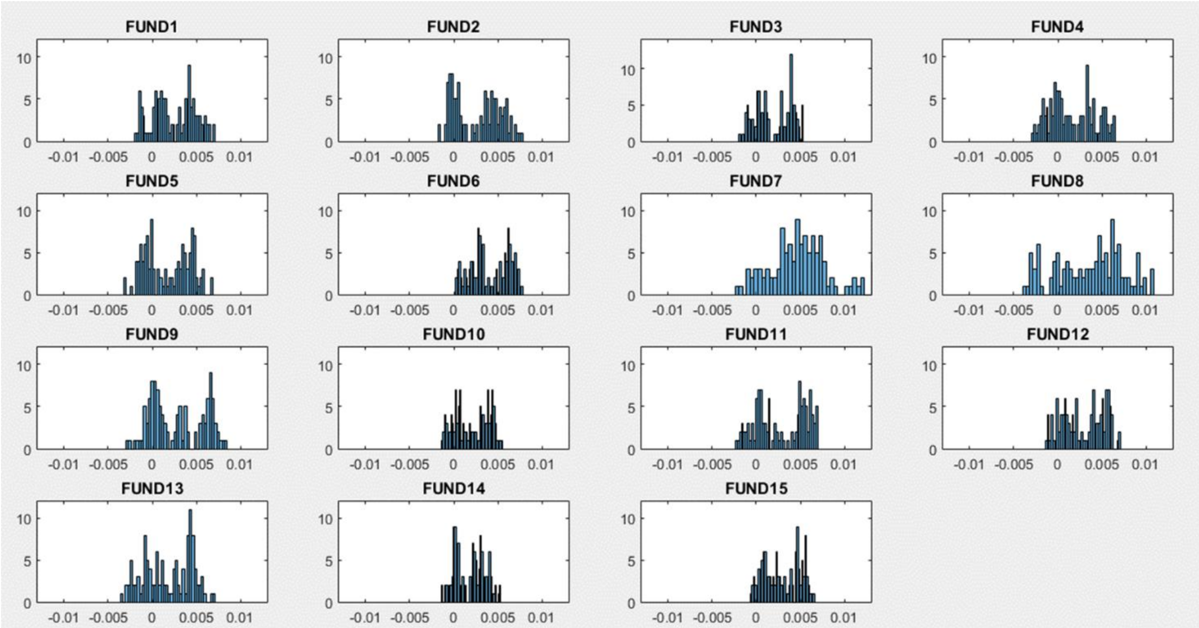



Source: own elaboration. Data obtained from Bloomberg Professional

The **Jarque-Bera's test** presented in the **Table 3** confirms this hypothesis. In fact, it rejects the null hypothesis of normality for all the funds at a 95% confidence level, apart for the highlighted FUND7.

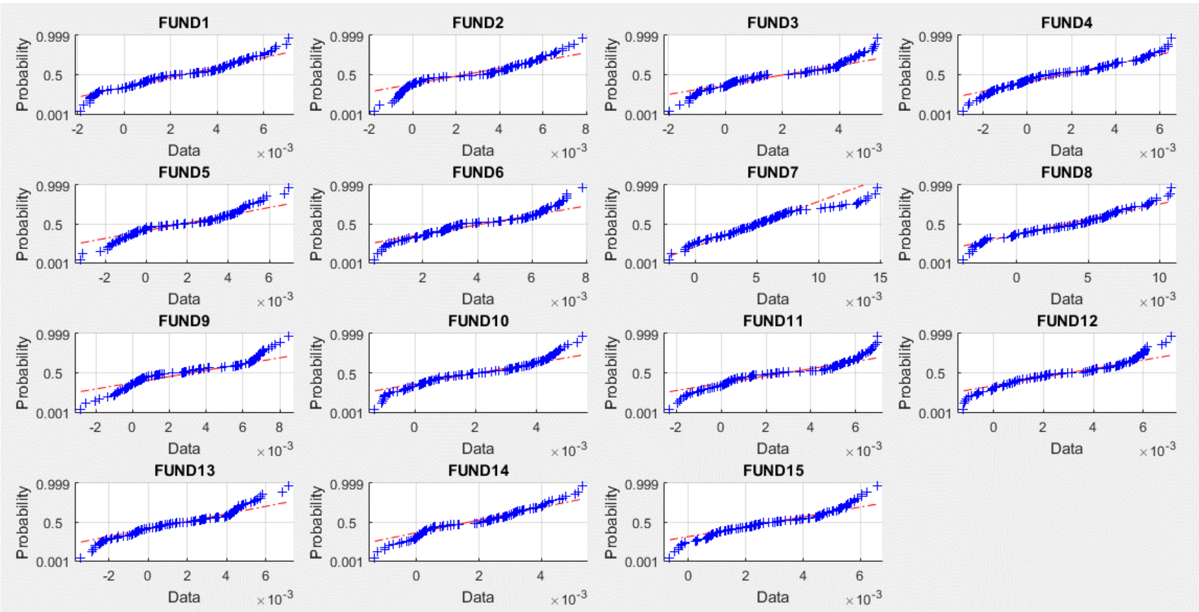
As regards to the 60 months rolling returns, the funds distributions do not show all the same characteristics. From Table 2 it can be seen that the skewness are both negative and positive, whereas the kurtosis are all negative, with the exception of the FUND7. However, once again, the **Jarque-Bera's test** rejects the null hypothesis of normality for all the funds at a 95% confidence level, except for one fund (FUND8). Below, we present the histograms **(Figure 8)** and the Q-Q plot **(Figure 9)** of month rolling distribution returns.

***Figure 8. Histograms of 60 months rolling returns***




Source: own elaboration. Data obtained from Bloomberg Professional

***Figure 9. Q-Q plots of 60 months rolling returns***




Source: own elaboration. Data obtained from Bloomberg Professional

In **summary**, excluding single cases (FUND 7 for monthly returns, FUND8 for 60 months rolling returns), the returns distributions of the analysed funds cannot be approximated by normal distributions.

# CHAPTER TWO

## 2 Performance measures: definitions and features

### 2.1 Introduction

In this chapter, we describe, from a theoretical point of view, the set of analysed **performance measures**.
The performance indicators can be grouped in several different and detailed ways. In this study, in order to maintain a simple structure, we divide these measures into **five groups** using the classification proposed by Caporin & Lisi (2011). The performance measures are classified from a statistical point of view, separating the use of general risk measures (group 1) from the ratios based on drawdowns (group 2), the ratios based on partial moments (group 3), the ratio based on quantiles (group 4) and the measures derived from utility functions (group 5).

In order to provide a general setup, we present **some adopted notation**:
$x_{i,t}$ is the price of asset i in period t;
$r_{i,t}$ is the return of asset i in period t;
$r_f$ is the return of risk free investment;
$r_{b,t}$ is the return of the benchmark in period t;
$\sigma\ [r_{i,t}]$ is the volatility of r;
$E\ [r^p]$ is the moment of order p of r;
$D^p_{i,t}$ is the $p^{th}$ larger drawdown of asset i in period t ($D^1_{i,t}$ is the max drawdown).

### 2.2 Traditional performance measures and other unclassified measures

The group of the **traditional performance measures and other unclassified measures** includes the Sharpe ratio, the Treynor index, the Jensen alpha, the information ratio, the $M^2$ measure, the adjusted for skewness and kurtosis Sharpe ratio and the Diaman ratio. This group thus contains the most known and traditional indices and other unclassified measures.

### 2.2.1 Sharpe ratio

In order to compare the risk to return efficiency of two funds, it is possible to divide the average return over a period by the risk taken during that period: this measure is called risk-adjusted return (Feibel, 2003).
William Sharpe, Nobel Prize winner and one of the originators of the Capital Asset Pricing Model, developed the **Sharpe ratio** for risk-adjusted performance measurement, that is a modification to the risk-adjusted return (Sharpe, 1966). The Sharpe ratio is the difference between the arithmetic mean of the fund returns and the risk-free return (called excess return), divided by the standard deviation of the fund returns.

$$\text{Sharpe ratio} = \frac{E[r_{i,t}] - r_f}{\sigma[r_{i,t}]} .$$

The modification of the Sharpe ratio compared to the risk-adjusted return has the effect of removing the portion of return brought by the risk-free rate, for which it is not expected to suffer any risk. The Sharpe ratio, revealing the risk/return efficiency of a portfolio, was called by the same William Sharpe reward to variability ratio. By using the Sharpe ratio, we do not compare only the absolute return earned on an investment but also the attended risk, in this case expressed by the proxy of the standard deviation. Higher the Sharpe ratio, higher the return provided by the fund per unit of risk.
The values of the Sharpe ratios differ depending on both the time and the type of investment analysed, thus it is difficult to offer an idea of what a good Sharpe ratio is. The key is to compare the Sharpe ratios with those of similar investments.

### 2.2.2 Treynor ratio

The **Treynor index** is calculated as a ratio between the average return of the portfolio above the risk-free rate and its systematic risk (Treynor, 1965). This ratio, having the Beta at the denominator, is drawn directly from the CAPM. It is expressed as follows:

$$\text{Treynor index} = \frac{E[r_{i,t}] - r_f}{\beta_i} .$$

In order to estimate the Beta of a portfolio we need to choose a reference index. The Treynor ratio is particularly appropriate for appreciating the performance of a well-diversified portfolio, since it only considers the **systematic risk** of a portfolio as a risk factor (Le Sourd, 2007). A part of the overall risk of a financial asset can be eliminated with diversification, the so-called diversifiable risk (or specific risk). Representing the particular risk of a specific asset, investors should not be remunerated to bear it; in fact, this risk can be avoided by diversification. On the contrary, a part of the overall risk cannot be eliminated, no matter how investors diversify their portfolio. This is the so-called systematic risk (or market risk, or not diversifiable), which can be considered as the risk of the market as a whole. The Beta express the exposure of a portfolio to the systematic risk.

### 2.2.3 Jensen's alpha

**Jensen's alpha** is defined as the differential between the excess return of a portfolio with regard to the risk-free rate and the return explained by the market model.

Jensen's alpha $= (E[r_{i,t}] - r_f) - \beta_i * (E[r_{b,t}] - r_f)$

The Jensen measure is based on the Capital Asset Pricing Model; in particular, it is the intercept of the regression equation in the CAPM ignoring the error term. If alpha is greater than zero, the fund have a higher return than the one expected by the CAPM (Jensen, 1968).
This difference is accordingly due to the active management of the portfolio manager. For this reason, this measure should evaluate the ability of a fund manager to select, for example, assets that are underpriced by the market.
This measure does not allow portfolios with different levels of risk to be compared because the value of alpha is proportional to the level of risk taken, measured by the Beta. Thus, the Jensen alpha should be used to rank portfolios within peer groups (Cogneau & Hubner, 2009).

### 2.2.4 Information ratio

The **information ratio** is built as a ratio between a numerator, defined as the excess return of a portfolio with regard to the benchmark, and a denominator, defined as the standard deviation of the difference between the returns of the portfolio and the returns of the benchmark (the so-called tracking error volatility) (Grinold & Kahn, 2000).

Information ratio $= \dfrac{E[r_{i,t}] - E[r_{b,t}]}{\sigma[r_{i,t} - r_{b,t}]}$

The information ratio is used to evaluate the active investment manager's skill compared to the benchmark. In particular, it is useful to estimate how the excess risk, taken as a consequence of an active strategy, is remunerated.
In other words, the information ratio states whether the fund manager is able to gain additional performance over the benchmark and whether he is able to obtain it without increasing the level of risk compared to the benchmark. Literally, the term information ratio refers to the idea that

the manager should deviate from the benchmark only if he has some special “information”, not already priced into the market, which presumably would lead him to add value over the benchmark return (Feibel, 2003).
Summarizing, the manager with the higher information ratio produces the higher excess return compared to the benchmark per unit of deviation from the benchmark.

### 2.2.5 M2 measure

Franco Modigliani, a Nobel Prize for economics, and Leah Modigliani, a Morgan Stanley analyst, are credited with developing the **$M^2$ measure** (Modigliani & Modigliani 1997). Their purpose was to help the investors to compare the returns that have been adjusted for risk.
The M2 measure is equivalent to the return that the fund would have achieved if it had had the same risk as the benchmark index.
It is expressed by the following formula:

$$M^2 = (E[r_{i,t}] - r_f) * \frac{\sigma[r_{b,t}]}{\sigma[r_{i,t}]} + r_f$$

According to Modigliani and Modigliani, this measure is easier to understand by the average investor than the Sharpe Ratio.
These two measure are directly proportional because the $M^2$ measure is nothing more than the Sharpe ratio scaled by the standard deviation of the benchmark return: for this reason, a relative ranking of funds created by the M2 measure will be coincident with the one created by the Sharpe ratio (Feibel, 2003). This rule holds if the analysed financial instruments has the same benchmark.
As the Sharpe ratio, the most interesting funds are those with the higher $M^2$ value.

### 2.2.6 Adjusted for skewness and kurtosis Sharpe ratio

The analysis of Gregoriou & Gueye (2003) shows how complications may occur when the traditional performance measures, as the Sharpe ratio, are used to evaluate funds characterized by distributions not normal. The traditional Sharpe ratio, for example, takes into account exclusively the first two moments of the distribution. However, as analysed in the previous section, the assumption of Gaussian distribution for return series does not hold for many funds. A statistical variation is proposed to tackle this issue, by including higher moments of the

distribution in the formula. In order to adjust for skewness and kurtosis, this index incorporates a penalty for excess kurtosis (positive or negative) and a penalty for negative skewness.
It is defined as follows:

$$\text{Adjusted Sharpe ratio} = \frac{E[r_{i,t}] - r_f}{\sigma[r_{i,t}]} - \left| Kurtosis(r_{i,t}) \right| + Skewness(r_{i,t}) .$$

### 2.2.7 Diaman ratio

The **Diaman ratio** was created in 2012 by Bernardi & Bertelli. While most of the performance indicators are constructed using the first two moments of the return distributions, the Diaman Ratio is a statistical indicator that does not need the returns to be normally distributed.
The Diaman ratio derives from the linear regression of a historical price series over time. It is expressed as the Beta multiplied by the coefficient of determination ($R^2$) (Bernardi & Bertelli, 2014).

$$\text{Diaman ratio} = \beta * R^2$$

Beta in this case is the slope of the regression line, whereas the $R^2$ is the coefficient of determination that defines the degree of goodness of fit. When the historical series tends to coincide with regression line, $R^2$ tends to one, whereas, when the historical series tends to deviate significantly from the regression line, it tends to zero.
The Diaman ratio is a statistical indicator that, without using the average and the standard deviation, should indicate the expected future return of a historical price series, corrected for the variability (expressed by the R^2).

## 2.3 Measures based on drawdown

These performance measures are **based on the concept of drawdown**. The drawdown is the measure of the decline from a peak. It is measured as the percentage between the peak and the subsequent trough (Burghardt & Walls, 2003). This indicator is very significant because it indicates, more clearly than the volatility, the real loss that an investor can suffer during the investment. Therefore, it allows the investor to reflect on whether such an investment represents a risk that he is willing to bear. This group of performance measure contains the following indices: Calmar ratio, Sterling ratio, Burke ratio and Ulcer index.

### 2.3.1 Calmar ratio

The idea of the **Calmar ratio** is to replace the standard deviation, used in the Sharpe ratio, by the maximum drawdown, a parameter that investors often consider.

$$\text{Calmar ratio} = \frac{E[r_{i,t}] - r_f}{|\,\mathrm{D}^1_{\mathrm{i,t}}\,|}$$

The maximum drawdown represents the maximum lost, in a considered period, compared to the value reached on a peak; it is no more than the largest, in absolute value, among the drawdowns. The Calmar ratio (Young, 1991) is simply the average return of an investment divided by its maximum loss, on the considered period. An obvious drawback of this measure is its sensitivity to outliers.

### 2.3.2 Sterling ratio

Precisely because of the sensitivity to outliers to which the Calmar ratio is subject, Sterling Jones proposed the Sterling ratio. However, a paper of this author describing this ratio was not found. For this reason, some people quoted Lars Kestner (1996) as the originator of this index, because he was the first who mentioned it in a paper.

In order to decrease the sensitivity to outliers, this ratio is constructed as following: the numerator is the arithmetic mean of the returns of the fund/asset, whereas the denominator is the average of the "w" largest drawdowns during the period. W is the parameter that identifies the number of values used in the calculation of the denominator of the index and we set it equal to 10 (Caporin & Lisi (2011) suggests to choose a w inside of the interval defined by [T/20-T/10], with T = length of the analysed sample).

$$\text{Sterling ratio} = \frac{E[r_{i,t}] - r_f}{|\,\frac{1}{w} * \sum_{j=1}^{w} D^{j}{}_{i,t}\,|}$$

Another version of the index suggests adding to the denominator an arbitrary threshold of 10%. It should adjust for the fact that short-term calculations of drawdown are understated (Cogneau & Hubner, 2009). However, given that the length of the analysed sample is quite long, it was used the version without the threshold.

### 2.3.3 Burke ratio

Similar to the Sterling ratio, the **Burke ratio** (Burke, 1994) discounts the expected excess return of the security by a factor derived by the worst "w" maximum drawdowns of the portfolio (w is set to 10). However, in this case, the denominator is the square root of the average of the squared "w" largest drawdowns.

Burke ratio $= \dfrac{E[r_{i,t}] - r_f}{(\frac{1}{w} * \sum_{j=1}^{w} [D^{j}{}_{i,t}]^2)^{\frac{1}{2}}}$

As the Sterling ratio, it is less sensitive to outliers.

### 2.3.4 Martin ratio

Martin & McCann (1989) propose a performance measure based on the Ulcer index, the **Martin ratio**. The numerator is the average excess return whereas the denominator is the Ulcer index, computed as the square root of the average of the squared drawdowns observed in that period. The Ulcer index measures the depth and the duration of percentage drawdowns.

Compared to the Shape ratio, the Martin ratio may present a concrete advantage: in the evaluation of the risk, it does not consider all the variability but only the downward changes (Cogneau & Hubner, 2009).

It is expressed as following:

Martin ratio $= \dfrac{E[r_{i,t}] - r_f}{(\frac{1}{T} * \sum_{j=1}^{T} [D_{i,t}]^2)^{\frac{1}{2}}}$.

## 2.4 Measures based on partial moments

At this point, we present the performance measures constructed using **partial moments** of the return distributions. The analysed indices are the following: Sortino index, K3 measure, Omega index and Farinelli-Tibiletti ratio.

### 2.4.1 Sortino index

Within this category, the most widely used measure is probably the **Sortino ratio**.

The generic expression of the Sortino ratio was formulated by Sortino & Satchell (2001). They developed a performance measure, called Reward to Lower Partial Moment ratio, based on the lower partial moments and formulated as follows:

RLPM $= \dfrac{E[r_{i,t}] - r_\tau}{(LPM_{r_i, r_\tau, o})^{\frac{1}{o}}}$, where $r_\tau$ is the generic threshold and "*o*" a positive constant.

In order to evaluate the performance of a fund, it sets the downside deviation as measure of risk. In the numerator, a generic threshold is subtracted to the average return (the same target return is considered in the computation of the semivariance at the denominator).
Sortino & Van Der Meer (1991) developed the **Sortino ratio**, characterized by "*o*" = 2 and a threshold equals to a minimum acceptable return.
It is defined as follows:

Sortino ratio $= \frac{E[r_{i,t}] - r_\tau}{E[(\min(r_{i,t} - r_\tau, 0))^2]^{\frac{1}{2}}}$, where $r_\tau$ is the target return.

We choose a target return equal to zero, corresponding to a neutral risk aversion.

### 2.4.2 K3 ratio

Kaplan & Knowles (2004) introduce a measure named Kappa3 ratio, or **K3 ratio**, which is comparable to the Sortino ratio except for the "*o*" equals to 3 and the absolute value at the denominator.

$$\text{K3 ratio } = \frac{E[r_{i,t}] - r_\tau}{|\, E[(\min(r_{i,t} - r_\tau, 0))^3]^{\frac{1}{3}} \,|}$$

### 2.4.3 Omega ratio

The **Omega ratio** is defined as follows (Menardi & Lisi, 2012a,b):

$$\text{Omega ratio} = \frac{E[r_{i,t}] - r_\tau}{|\, E[\min(r_{i,t} - r_\tau, 0)] \,|} + 1 .$$

### 2.4.4 Farinelli-Tibiletti ratio

**Farinelli & Tibiletti** (2003) propose a generalized performance measure defined as the ratio between an upper partial moment of order p and a lower partial moment of order q.

$$\text{FT ratio } = \frac{E[\left|\max(r_{i,t} - r_\tau, 0)\right|^p]^{\frac{1}{p}}}{E[\left|\min(r_{i,t} - r_\tau, 0)\right|^q]^{\frac{1}{q}}}$$

Two types of asymmetric preferences can be modelled using this indicator: 1) the asymmetric preference between “good” and “bad” volatility from the benchmark; 2) the asymmetric preference between small and large deviations from the benchmark.

The values assigned to p and q are used to form the indicator: higher p and q, higher the investor’s preference for (expected gains with p) or dislike of (expected losses for q) extreme events (Cogneau & Hubner, 2009).

Thus, the partial moment orders p and q are calibrated in order to match them with possible investors’ preferences: p=0,5 and q=2 for a defensive investor; p=1,5 and q=2 for a conservative investor; p=2 and q=1,5 for a “growth” investor; p=3 and q=0,5 for an aggressive investor (Farinelli et al., 2009).

## 2.5 Measures based on quantiles

This group of performance **measures is based on quantiles** of return distributions. We firstly need to define the following relevant quantities: the **Value at Risk** and the **Conditional Value at Risk** (or expected shortfall).

J.P. Morgan is credited with helping to make **VaR** a widely used measure from the 90s (Hull, 2012). It indicates the worst loss an investor is expected to suffer at a certain confidence level in a certain period of time. Statistically it is the alpha-quantile, where alpha is defined as one minus the confidence level, of the return distribution.

$P[r_{i,t} \leq VaR(r_{i,t}, 1-\alpha)] = \alpha$, where (1-α) is the confidence level.

However, the Value at Risk does not incorporate in its value what happens in the left tail.

The **Conditional Value at Risk**, or expected shortfall, answer the question: “If the investment goes wrong, which is the expected loss?”.

It is defined as the expected return of the returns smaller than the VaR.

$$ES[r_{i,t}, 1-\alpha] = E[r_{i,t} \mid r_{i,t} \leq VaR(r_{i,t}, 1-\alpha)]$$

For all the indicators belonging to the category of the measures based on quantiles we set the alpha equal to 5%, corresponding to a 95% level of confidence.

### 2.5.1 VR ratio

The **VR ratio** is defined as the expected excess return over the absolute value of VaR at a given confidence level (Caporin & Lisi, 2011). The Value at Risk is thus the measure of risk of this index.

$$\text{VR} = \frac{E[r_{i,t}] - r_f}{|VaR(r_{i,t}, 95\%)|}$$

### 2.5.2 VARR ratio

The **VARR ratio** (Caporin & Lisi, 2011) is defined as the ratio between the absolute value of VaR of opposite returns and the absolute value of VaR of "intact" returns. It is basically the absolute value of the 95$^{th}$ quantile of the return distribution over the absolute value of the 5$^{th}$ quantile of the return distribution.

$$\text{VARR} = \frac{|VaR(-r_{i,t}, 95\%)|}{|VaR(r_{i,t}, 95\%)|}$$

### 2.5.3 STARR ratio

The **STARR (Stable Tail Adjusted Return Ratio) ratio** is defined as the expected return over the absolute value of the Conditional Value at Risk (Martin, Rachev & Siboulet, 2003).

$$\text{STARR} = \frac{E[r_{i,t}] - r_f}{|CVaR(r_{i,t}, 95\%)|}$$

### 2.5.4 Generalized Rachev ratio

Biglova et al. proposed the **generalized Rachev ratio** in 2004. It is defined as follows:

$$\text{GRR ratio} = \frac{E[|r_{i,t}|^p \,|\, r_{i,t} \geq -VaR(-r_{i,t}, 95\%)]^{\frac{1}{p}}}{E[|r_{i,t}|^q \,|\, r_{i,t} \leq VaR(r_{i,t}, 95\%)]^{\frac{1}{q}}}$$

The numerator is the "p-root" of the absolute value of the expected value of the returns greater than the value of the 95$^{th}$ percentile powered to "p". The denominator is the "q-root" of the absolute value of expected value of the returns smaller than the value of the 5$^{th}$ percentile elevated to "q". The power indices vary according to the investor's degree of risk aversion and attraction to high returns. We set three different cases: the combination p=0,5 and q=2 for a risk averse investor, the combination p=q=1 (that gives the simple Rachev Ratio) for a risk neutral investor and the combination p=3 and q=0,5 for a risk lover investor.

## 2.6 Measures derived from utility functions

The indicators belonging to this group **derived from utility functions**. These performance measures, expressed per unit of marginal utility, incorporate the investors' preferences and risk profiles through representative utility functions. This category is represented by the Morningstar risk-adjusted return.

### 2.6.1 Morningstar risk adjusted return

The original Morningstar Rating was introduced in 1985; since its creation, it helps the investors selecting the funds in which to invest among those available.
The methodology (Morningstar, 2007) that allows Morningstar to publish this ranking of funds is based on funds' risk-adjusted returns. Applying expected utility theory to risk-adjusted return**,** the measure quantify how investors feel about one distribution of returns versus another. Therefore, it estimates the utility provided by a fund to an investor that has a power utility function that depends on the value of the risk-aversion coefficient (Lisi & Caporin, 2012). Theoretically, this coefficient can assume any value, without constraints. When it is less then (-1), the investor is risk-loving. When it is equal to (-1), the degree of risk aversion is zero. When is greater than zero, the investor demands a larger risk premium for choosing the risky portfolio. The larger the coefficient, the more risk averse the investor (Lisi and Caporin, 2012). We set three different parameters for the risk-aversion coefficient, indicated as lambda: 2, 20, 50.
In case of monthly returns, the Morningstar risk-adjusted return is defined as follows:

$$\text{MRAR} = \begin{cases} E[(1+r_{i,t})^{-\lambda}]^{-\frac{12}{\lambda}} & \lambda > -1, \lambda \neq 0 \\ e^{E[\ln(1+r_{i,t})]} & \lambda = 0 \end{cases}$$

Despite its widespread use, the measure refers to a power utility function displaying an unrealistic constant relative risk aversion coefficient over time.

## 2.7 Conclusions

The following table **(Table 5)** summarizes the described performance measures.

***Table 5. Performance measures included in the analysis***

| | |
|---|---|
| **Traditional Performance Measures and Other Unclassified Measures** | |
| **SHARPE RATIO** | $\frac{E[r_{i,t}]-r_f}{\sigma[r_{i,t}]}$ |
| **TREYNOR INDEX** | $\frac{E[r_{i,t}]-r_f}{\beta_i}$ |
| **JENSEN'S ALPHA** | $(E[r_{i,t}]-r_f)-\beta_i*(E[r_{b,t}]-r_f)$ |
| **INFORMATION RATIO** | $\frac{E[r_{i,t}]-E[r_{b,t}]}{\sigma[r_{i,t}-r_{b,t}]}$ |
| **M$^2$ MEASURE** | $(E[r_{i,t}]-r_f)*\frac{\sigma[r_{i,t}]}{\sigma[r_{b,t}]}+r_f$ |
| **ADJUSTED FOR SKEWNESS AND KURTOSIS SHARPE RATIO** | $\frac{E[r_{i,t}]-r_f}{\sigma[r_{i,t}]}-\lvert Kurtosis(r_{i,t})\rvert+Skewness(r_{i,t})$ |
| **DIAMAN RATIO** | $\beta*R^2$ |
| **Measures based on Drawdown** | |
| **CALMAR RATIO** | $\frac{E[r_{i,t}]-r_f}{\lvert \mathrm{D}^1{}_{\mathrm{i,t}} \rvert}$ |
| **STERLING RATIO** | $\frac{E[r_{i,t}]-r_f}{\lvert \frac{1}{w}*\sum_{j=1}^{w} D^j{}_{i,t} \rvert}$ |
| **BURKE RATIO** | $\frac{E[r_{i,t}]-r_f}{(\frac{1}{w}*\sum_{j=1}^{w}[D^j{}_{i,t}]^2)^{\frac{1}{2}}}$ |
| **MARTIN RATIO** | $\frac{E[r_{i,t}]-r_f}{(\frac{1}{T}*\sum_{j=1}^{T}[D_{i,t}]^2)^{\frac{1}{2}}}$ |
| **Measures based on Partial Moments** | |
| **SORTINO INDEX** | $\frac{E[r_{i,t}]-r_\tau}{E[(\min(r_{i,t}-r_\tau,0))^2]^{\frac{1}{2}}}$ |
| **K3 RATIO** | $\frac{E[r_{i,t}]-r_\tau}{\lvert E[(\min(r_{i,t}-r_\tau,0))^3]^{\frac{1}{3}} \rvert}$ |
| **OMEGA RATIO** | $\frac{E[r_{i,t}]-r_\tau}{\lvert E[\min(r_{i,t}-r_\tau,0)]\rvert}+1$ |
| **FARINELLI TIBILETTI RATIO** | $\frac{E[\lvert\max(r_{i,t}-r_\tau,0)\rvert^p]^{\frac{1}{p}}}{E[\lvert\min(r_{i,t}-r_\tau,0)\rvert^q]^{\frac{1}{q}}}$ |
| **Measures Based on Quantiles** | |
| **VR RATIO** | $\frac{E[r_{i,t}]-r_f}{\lvert VaR(r_{i,t},95\%)\rvert}$ |
| **VARR RATIO** | $\frac{\lvert VaR(-r_{i,t},95\%)\rvert}{\lvert VaR(r_{i,t},95\%)\rvert}$ |
| **STARR RATIO** | $\frac{E[r_{i,t}]-r_f}{\lvert CVaR(r_{i,t},95\%)\rvert}$ |
| **GENERALIZED RACHEV RATIO** | $\frac{E[\lvert r_{i,t}\rvert^p \mid r_{i,t}\geq -VaR(-r_{i,t},95\%)]^{\frac{1}{p}}}{E[\lvert r_{i,t}\rvert^q \mid r_{i,t}\leq VaR(r_{i,t},95\%)]^{\frac{1}{q}}}$ |
| **Measures Derived from Utility Functions** | |
| **MORNINGSTAR RISK ADJUSTED RETURN** | $\begin{cases} E[(1+r_{i,t})^{-\lambda}]^{-\frac{12}{\lambda}} & \lambda>-1,\lambda\neq 0 \\ e^{E[\ln(1+r_{i,t})]} & \lambda=0 \end{cases}$ |

Source: own elaboration

CHAPTER THREE

# 3 Analysis of correlation, stability and other features of examined performance measures

## 3.1 Introduction

This chapter is related to the analysis of some features (in particular, correlation and stability) of the presented performance measures.

## 3.2 Correlation between performance measures

The first section of this chapter regards an analysis of the possible **mutual correlation** between the performance measures. During this process, we firstly perform a **static analysis** of the rank correlation. Secondly, we use a **rolling approach** in order to test the correlation between the rankings over time. Based on this correlation analysis, we finally select a group of uncorrelated performance measures.

### 3.2.1 Introduction

When some performance measures are highly correlated, they may be considered **redundant**. In fact, in that case, it means that they bring the same information and some of them may be abandoned. We present the correlation analysis in order to select a limited group of performance measure carrying **different information**, exploiting the possible mutual correlations between the indicators.

In fact, an enormous number of performance measures are published in the academic literature aimed to measure or analyse portfolios performances. As stated in the previous chapter, in addition to the classical indicators, alternative measures are more and more frequently taken into consideration. They are mainly constructed in order to satisfy some particular necessity or to overcome the limits of some traditional measure (for example, the assumption of normality for the return distribution). The aim of these indicators is to try to bring some new information compared to the classical ones; however, they may implicitly provide the same knowledge.

We develop this correlation analysis in order to reduce the number of performance measures, taking into account just those that really carry different information.

Some authors has already presented a comparison between indicators using their mutual correlation. Among all, Gemmill et al. (2006), Eling & Schuhmacher (2007), Eling et al. (2011) and Caporin & Lisi (2011). In this study we follow the methodologies adopted by Eling and Schuhmacher (2007) and Caporin & Lisi (2011), exploiting the information provided by the **rank correlations**.

In particular, we base our analysis on mutual funds as in Eling and Schuhmacher (2007). According to Caporin & Lisi (2011), we set up a decision rule to define when two performance measures are "highly correlated". However, we introduce some additional **extensions**. First, we analyse some new performance measure, such as the Diaman ratio, the adjusted for skewness and kurtosis Sharpe ratio and the Martin ratio. Moreover, in contrast to the previous studies, we analyse the correlations between the performance measures **belonging to the same group** (in Caporin & Lisi, for example, the first selection of performance measures is conducted through an analysis of the correlation between indicators that differ only for the parameters included in their definition).

We thus analyse the possible relationships between the performance measures through the dynamic evolution of the **fund rankings** induced by the various indicators. Different ranks are the outcome of different informative content of the performance measures which created those classifications.

### 3.2.2 Theory and methodology used to study the correlation

We examine the correlation between rankings using the **Spearman rank correlation** ($\rho_s$).

We calculate all the performance measures empirically, applying the real sample moments and sample quantiles.

The aim is to analyse the **degree of correlation** between the rankings induced by the different indicators, detecting a level of correlation above which two measures can be considered highly correlated.

In order to define this threshold we define as "low" a correlation smaller than 0,8 ($\rho_s \leq 0{,}8$). We then consider the asymptotic distribution of $\rho_s$, in order to have a precise threshold.

N is defined as the number of analysed financial assets and $z$ is the Fisher transformation of $\rho_s$.

In particular,

$z := \frac{1}{2} \ln \frac{(1+\hat{\rho}_s)}{(1-\hat{\rho}_s)}$, with $\hat{\rho}_s$ and $\hat{z}$ corresponding to the sample quantities.

Asymptotically,

$\sqrt{N-2\hat{\rho}} \approx \mathrm{N}ormal(\rho,1)$ (Caporin & Lisi, 2011).

The threshold is thus defined as:

$$\rho^*{}_s(\alpha) = \frac{e^{\ln(\frac{1+\rho_s}{1-\rho_s})+2Z_{1-\alpha}\sqrt{\frac{1}{N-2}}} - 1}{e^{\ln(\frac{1+\rho_s}{1-\rho_s})+2Z_{1-\alpha}\sqrt{\frac{1}{N-2}}} + 1}.$$

It corresponds to the critical value, where α is the significance level and $Z_{1-\alpha}$ is the $(1-\alpha)^{th}$ quantile of a standard normal distribution.

With N = 15 and α = 5%, the critical value $\rho^*_s$ defining a **low correlation** turns out to be **0,915**.

As previously specified, the threshold level depends on the sample dimension. For a small number of assets, as in our study, the critical value results quite large, easily leading to an acceptance of the null hypothesis of independence. We will take into account this **"bias"** in the rest of the analysis.

### 3.2.3 Static analysis of the rank correlation

In this section, we report the **static analysis of the rank correlation,** implemented using four different **evaluation windows**.

We conduct the analysis only on the simple returns because, as demonstrated in Caporin & Lisi (2011), the use of alternative return types (for example, excess return with respect to the benchmark or the risk free rate) does not affect the evaluation of rank correlations.

At first, the focus is on the entire sample, equivalent to 178 monthly returns. The other windows are built on reduced set of data. In particular, the second one is composed by the last 120 returns (from March 2007 to February 2017), the third one by the last 60 returns (from March 2012 to February 2017) whereas the last one by the last 36 returns (from March 2014 to February 2017).

For every time window, we evaluate if the performance measures are correlated. We apply the performance measures on all the returns that compose a window and we calculate the Spearman rank correlation on the resulting indicators.

The correlations between measures will be analysed **within** their **peer group**, following the classification previously described.

In the next **Tables,** we show the average Spearman rank correlations calculated for each pair of indicators belonging to the same category. For each group, the first table refers to the entire sample, the second one to the sample of the last 120 returns, the third one to the last 60 returns and the fourth to the last 36 returns.

The **highlighted values** identify rank correlations below the threshold level of 0,915, recognizing the measures for which we could accept the hypothesis of **independence**. On the contrary, if the correlations are greater than 0,915, the measures are "highly correlated" and one of them may potentially be excluded from the forthcoming analysis; carrying the same informative content, there is no need to consider both the indicators.

### 3.2.3.1 Traditional and other unclassified performance measures

The first analysed group is relative to the **traditional and other unclassified performance measures (Table 6**, **Table 7**, **Table 8** and **Table 9)**.

***Table 6. Rank correlations across selected performances measures - traditional and other unclassified performance measures. 178 monthly returns window***

| 178 RETURNS | SHARPE | TREYNOR | ALPHA J | INF. RATIO | $M^2$ | ADJ. SHARPE | DIAMAN |
|---|---|---|---|---|---|---|---|
| SHARPE | 1 | 0,850 | 0,825 | 0,650 | 1,000 | 0,089 | 0,596 |
| TREYNOR | 0,850 | 1 | 0,954 | 0,800 | 0,850 | 0,093 | 0,679 |
| ALPHA J | 0,825 | 0,954 | 1 | 0,650 | 0,825 | 0,157 | 0,529 |
| INF. RATIO | 0,650 | 0,800 | 0,650 | 1 | 0,650 | 0,032 | 0,736 |
| $M^2$ | 1,000 | 0,850 | 0,825 | 0,650 | 1 | 0,089 | 0,596 |
| ADJ. SHARPE | 0,089 | 0,093 | 0,157 | 0,032 | 0,089 | 1 | 0,089 |
| DIAMAN | 0,596 | 0,679 | 0,529 | 0,736 | 0,596 | 0,089 | 1 |

Source: own elaboration

***Table 7. Rank correlations across selected performances measures - traditional and other unclassified performance measures. 120 monthly returns window***

| 120 RETURNS | SHARPE | TREYNOR | ALPHA J | INF. RATIO | $M^2$ | ADJ. SHARPE | DIAMAN |
|---|---|---|---|---|---|---|---|
| SHARPE | 1 | 0,964 | 0,911 | 0,375 | 1,000 | 0,468 | 0,879 |
| TREYNOR | 0,964 | 1 | 0,911 | 0,500 | 0,964 | 0,596 | 0,846 |
| ALPHA J | 0,911 | 0,911 | 1 | 0,379 | 0,911 | 0,536 | 0,768 |
| INF. RATIO | 0,375 | 0,500 | 0,379 | 1 | 0,375 | 0,464 | 0,393 |
| $M^2$ | 1,000 | 0,964 | 0,911 | 0,375 | 1 | 0,468 | 0,879 |
| ADJ. SHARPE | 0,468 | 0,596 | 0,536 | 0,464 | 0,468 | 1 | 0,500 |
| DIAMAN | 0,879 | 0,846 | 0,768 | 0,393 | 0,879 | 0,500 | 1 |

Source: own elaboration

***Table 8. Rank correlations across selected performances measures - traditional and other unclassified performance measures. 60 monthly returns window***

| 60 RETURNS | SHARPE | TREYNOR | ALPHA J | INF. RATIO | $M^2$ | ADJ. SHARPE | DIAMAN |
|---|---|---|---|---|---|---|---|
| SHARPE | 1 | 0,814 | 0,739 | 0,314 | 1,000 | -0,336 | 0,664 |
| TREYNOR | 0,814 | 1 | 0,954 | 0,543 | 0,814 | 0,007 | 0,857 |
| ALPHA J | 0,739 | 0,954 | 1 | 0,396 | 0,739 | -0,043 | 0,732 |
| INF. RATIO | 0,314 | 0,543 | 0,396 | 1 | 0,314 | 0,507 | 0,800 |
| $M^2$ | 1,000 | 0,814 | 0,739 | 0,314 | 1 | -0,336 | 0,664 |
| ADJ. SHARPE | -0,336 | 0,007 | -0,043 | 0,507 | -0,336 | 1 | 0,168 |
| DIAMAN | 0,664 | 0,857 | 0,732 | 0,800 | 0,664 | 0,168 | 1 |

Source: own elaboration

***Table 9. Rank correlations across selected performances measures - traditional and other unclassified performance measures. 36 monthly returns window***

| **36 RETURNS** | SHARPE | TREYNOR | ALPHA J | INF. RATIO | $M^2$ | ADJ. SHARPE | DIAMAN |
|---|---|---|---|---|---|---|---|
| SHARPE | **1** | **0,936** | **0,768** | **0,246** | **1,000** | **-0,107** | **0,650** |
| TREYNOR | **0,936** | **1** | **0,579** | **0,225** | **0,936** | **0,014** | **0,514** |
| ALPHA J | **0,768** | **0,579** | **1** | **0,407** | **0,768** | **-0,143** | **0,764** |
| INF. RATIO | **0,246** | **0,225** | **0,407** | **1** | **0,246** | **0,354** | **0,754** |
| $M^2$ | **1,000** | **0,936** | **0,768** | **0,246** | **1** | **-0,107** | **0,650** |
| ADJ. SHARPE | **-0,107** | **0,014** | **-0,143** | **0,354** | **-0,107** | **1** | **0,204** |
| DIAMAN | **0,650** | **0,514** | **0,764** | **0,754** | **0,650** | **0,204** | **1** |

Source: own elaboration

The first aspect that we can note is that, as previously described, the **Sharpe** ratio and the **$M^2$ measure** are perfectly correlated: their ranking correlations are always equal to one. As analysed in the previous chapter, these two measure are directly proportional; M2 measure is no more than the Sharpe ratio scaled by a factor. For this reason, we can consider just one of these two indicators without losing potentially useful information.

The **information ratio**, the **adjusted Sharpe ratio** and the **Diaman ratio** are slightly correlated to the other indicators of the group, for all the time windows. In order not to lose possible valuable information, we have to take into consideration all these measures for the rest of the analysis.

The rank correlation between **Sharpe ratio** and **Jensen's alpha** is under the threshold value in every window. We thus have to consider both the measures for the rest of the analysis.

Regarding the **Treynor ratio**, it provides rankings highly correlated with both **Sharpe ratio** and **Jensen's alpha**, but not in all the evaluation windows. It is highly correlated with the Sharpe ratio in the windows of 120 and 36 returns, and with the Jensen's alpha in windows of 176 and 60 returns. For this reason, we will study these possible relationships more in details in the section referred to the rolling analysis.

### 3.2.3.2 Measures based on drawdown

At this point, we examine the group of **measures based on drawdown**.

***Table 10. Rank correlations across selected performances measures - measures based on Drawdown. 178 monthly returns window***

| **178 RETURNS** | CALMAR | STERLING | BURKE | MARTIN |
|---|---|---|---|---|
| CALMAR | **1** | **0,989** | **0,989** | **0,961** |
| STERLING | **0,989** | **1** | **1,000** | **0,957** |
| BURKE | **0,989** | **1,000** | **1** | **0,957** |
| MARTIN | **0,961** | **0,957** | **0,957** | **1** |

Source: own elaboration

***Table 11. Rank correlations across selected performances measures - measures based on Drawdown. 120 monthly returns window***

| **120 RETURNS** | CALMAR | STERLING | BURKE | MARTIN |
|---|---|---|---|---|
| CALMAR | **1** | **0,996** | **0,996** | **0,982** |
| STERLING | **0,996** | **1** | **1,000** | **0,989** |
| BURKE | **0,996** | **1,000** | **1** | **0,989** |
| MARTIN | **0,982** | **0,989** | **0,989** | **1** |

Source: own elaboration

***Table 12. Rank correlations across selected performances measures - measures based on Drawdown. 60 monthly returns window***

| **60 RETURNS** | CALMAR | STERLING | BURKE | MARTIN |
|---|---|---|---|---|
| CALMAR | **1** | **0,939** | **0,943** | **0,907** |
| STERLING | **0,939** | **1** | **0,996** | **0,979** |
| BURKE | **0,943** | **0,996** | **1** | **0,986** |
| MARTIN | **0,907** | **0,979** | **0,986** | **1** |

Source: own elaboration

***Table 13. Rank correlations across selected performances measures - measures based on Drawdown. 36 monthly returns window***

| **36 RETURNS** | CALMAR | STERLING | BURKE | MARTIN |
|---|---|---|---|---|
| CALMAR | **1** | **0,921** | **0,936** | **0,904** |
| STERLING | **0,921** | **1** | **0,996** | **0,996** |
| BURKE | **0,936** | **0,996** | **1** | **0,989** |
| MARTIN | **0,904** | **0,996** | **0,989** | **1** |

Source: own elaboration

The indicators belonging to this group of measures show a high **within group** rank correlation. We find cases of low rank correlation just in a few occasions. In particular, when the window is formed by 36 returns, the **Calmar ratio** exhibit a low correlation with both **Martin ratio** and **Sterling ratio**. When the window is formed by 60 returns, the **Calmar ratio** is slightly correlation only with **Martin ratio**.

In order to detect if we can choose one indicator as representative for the entire group, we will examine these two relationships through the rolling analysis.

### 3.2.3.3 Measures based on partial moments

The following tables are referred to the **measures based on partial moments**.

A first evidence attests that a sub-group of measures, composed by the **Sortino ratio**, the **K3 ratio** and the **Omega ratio**, is characterized by a high within group correlation (except for the relationship between Omega and K3 in one window, case that will be examined through the rolling approach).

The **Farinelli-Tibiletti for a defensive investor** is highly correlated with both **Sortino ratio** and **K3 ratio**, excluding the window composed by 120 returns. It may be reasonable to choose just one index among this sub-group.

The other sub-group is composed by the **Farinelli-Tibiletti for a conservative investor**, the **Farinelli-Tibiletti for a "growth" investor** and the **Farinelli-Tibiletti for an aggressive investor**. The last measure shows a low correlation with all the other indices of the group; thus, we cannot exclude it from the rest of analysis. The Farinelli-Tibiletti for a conservative investor and the Farinelli-Tibiletti for a "growth" investor are highly correlated in three out of the four windows. If the rolling analysis will confirm this relationship, we may select just one of these two indicator without losing information.

***Table 14*. Rank correlations across selected performances measures - measures based on partial moments. 178 monthly returns window**

| 178 RETURNS | SORTINO | K3 | OMEGA | FT DEF. | FT CONS. | FT GROWTH | FT AGGR. |
|---|---|---|---|---|---|---|---|
| SORTINO | 1 | 0,975 | 0,954 | 0,925 | 0,779 | 0,761 | 0,675 |
| K3 | 0,975 | 1 | 0,914 | 0,936 | 0,821 | 0,804 | 0,682 |
| OMEGA | 0,954 | 0,914 | 1 | 0,814 | 0,639 | 0,682 | 0,761 |
| FT DEF. | 0,925 | 0,936 | 0,814 | 1 | 0,836 | 0,725 | 0,557 |
| FT CONS. | 0,779 | 0,821 | 0,639 | 0,836 | 1 | 0,907 | 0,371 |
| FT GROWTH | 0,761 | 0,804 | 0,682 | 0,725 | 0,907 | 1 | 0,568 |
| FT AGGR. | 0,675 | 0,682 | 0,761 | 0,557 | 0,371 | 0,568 | 1 |

Source: own elaboration

***Table 15. Rank correlations across selected performances measures - measures based on partial moments. 120 monthly returns window***

| 120 RETURNS | SORTINO | K3 | OMEGA | FT DEF. | FT CONS. | FT GROWTH | FT AGGR. |
|---|---|---|---|---|---|---|---|
| SORTINO | 1 | 0,989 | 0,979 | 0,879 | 0,739 | 0,764 | 0,589 |
| K3 | 0,989 | 1 | 0,954 | 0,889 | 0,804 | 0,818 | 0,561 |
| OMEGA | 0,979 | 0,954 | 1 | 0,857 | 0,639 | 0,679 | 0,657 |
| FT DEF. | 0,879 | 0,889 | 0,857 | 1 | 0,714 | 0,621 | 0,396 |
| FT CONS. | 0,739 | 0,804 | 0,639 | 0,714 | 1 | 0,939 | 0,236 |
| FT GROWTH | 0,764 | 0,818 | 0,679 | 0,621 | 0,939 | 1 | 0,411 |
| FT AGGR. | 0,589 | 0,561 | 0,657 | 0,396 | 0,236 | 0,411 | 1 |

Source: own elaboration

***Table 16. Rank correlations across selected performances measures - measures based on partial moments. 60 monthly returns window***

| 60 RETURNS | SORTINO | K3 | OMEGA | FT DEF. | FT CONS. | FT GROWTH | FT AGGR. |
|---|---|---|---|---|---|---|---|
| SORTINO | 1 | 0,989 | 0,971 | 0,989 | 0,889 | 0,904 | 0,821 |
| K3 | 0,989 | 1 | 0,954 | 0,986 | 0,914 | 0,911 | 0,771 |
| OMEGA | 0,971 | 0,954 | 1 | 0,975 | 0,804 | 0,846 | 0,900 |
| FT DEF. | 0,989 | 0,986 | 0,975 | 1 | 0,875 | 0,879 | 0,818 |
| FT CONS. | 0,889 | 0,914 | 0,804 | 0,875 | 1 | 0,971 | 0,661 |
| FT GROWTH | 0,904 | 0,911 | 0,846 | 0,879 | 0,971 | 1 | 0,768 |
| FT AGGR. | 0,821 | 0,771 | 0,900 | 0,818 | 0,661 | 0,768 | 1 |

Source: own elaboration

***Table 17. Rank correlations across selected performances measures - measures based on partial moments. 36 monthly returns window***

| 36 RETURNS | SORTINO | K3 | OMEGA | FT DEF. | FT CONS. | FT GROWTH | FT AGGR. |
|---|---|---|---|---|---|---|---|
| SORTINO | 1 | 0,986 | 0,957 | 0,936 | 0,721 | 0,839 | 0,721 |
| K3 | 0,986 | 1 | 0,918 | 0,925 | 0,814 | 0,896 | 0,632 |
| OMEGA | 0,957 | 0,918 | 1 | 0,879 | 0,561 | 0,725 | 0,854 |
| FT DEF. | 0,936 | 0,925 | 0,879 | 1 | 0,682 | 0,746 | 0,646 |
| FT CONS. | 0,721 | 0,814 | 0,561 | 0,682 | 1 | 0,950 | 0,157 |
| FT GROWTH | 0,839 | 0,896 | 0,725 | 0,746 | 0,950 | 1 | 0,357 |
| FT AGGR. | 0,721 | 0,632 | 0,854 | 0,646 | 0,157 | 0,357 | 1 |

Source: own elaboration

### 3.2.3.4 Measures based on quantiles

The next group is relative to the **measures based on quantiles**.

We found only low correlations between the **VARR ratio** and the other measures of the group; consequently, we have to select it for the rest of the study.

The **VR ratio** and the **STARR ratio** seem to provide highly correlated rankings to each other, except for the 60 returns window (their relationship will be studied through the rolling analysis). On the contrary, they show low correlations with the other indices of the group. One of the two measures has to be taken into consideration for the rest of the analysis.

The sub-group composed by the three **GRR ratios** is characterized by a high within group rank correlation. Even though in Caporin & Lisi (2011) the GRR for risk lover investors shows low correlations to the other measures, given our results we can undoubtedly consider just one of these three indices.

***Table 18. Rank correlations across selected performances measures - measures based on quantiles. 178 monthly returns window***

| 178 RETURNS | VR | VARR | STARR | GRR RISK NEUT. | GRR RISK AV. | GRR RISK LOV. |
|---|---|---|---|---|---|---|
| VR | 1 | 0,661 | 0,946 | 0,479 | 0,436 | 0,400 |
| VARR | 0,661 | 1 | 0,654 | 0,489 | 0,486 | 0,339 |
| STARR | 0,946 | 0,654 | 1 | 0,589 | 0,564 | 0,479 |
| GRR RISK NEUT. | 0,479 | 0,489 | 0,589 | 1 | 0,993 | 0,943 |
| GRR RISK AV. | 0,436 | 0,486 | 0,564 | 0,993 | 1 | 0,925 |
| GRR RISK LOV. | 0,400 | 0,339 | 0,479 | 0,943 | 0,925 | 1 |

Source: own elaboration

***Table 19. Rank correlations across selected performances measures - measures based on quantiles. 120 monthly returns window***

| 120 RETURNS | VR | VARR | STARR | GRR RISK NEUT. | GRR RISK AV. | GRR RISK LOV. |
|---|---|---|---|---|---|---|
| VR | 1 | 0,482 | 0,961 | 0,518 | 0,525 | 0,354 |
| VARR | 0,482 | 1 | 0,543 | 0,671 | 0,636 | 0,468 |
| STARR | 0,961 | 0,543 | 1 | 0,661 | 0,668 | 0,507 |
| GRR RISK NEUT. | 0,518 | 0,671 | 0,661 | 1 | 0,996 | 0,950 |
| GRR RISK AV. | 0,525 | 0,636 | 0,668 | 0,996 | 1 | 0,954 |
| GRR RISK LOV. | 0,354 | 0,468 | 0,507 | 0,950 | 0,954 | 1 |

Source: own elaboration

***Table 20. Rank correlations across selected performances measures - measures based on quantiles. 60 monthly returns window***

| **60 RETURNS** | VR | VARR | STARR | GRR RISK NEUT. | GRR RISK AV. | GRR RISK LOV. |
|---|---|---|---|---|---|---|
| VR | **1** | **0,596** | **0,871** | **0,246** | **0,179** | **0,279** |
| VARR | **0,596** | **1** | **0,546** | **0,311** | **0,264** | **0,339** |
| STARR | **0,871** | **0,546** | **1** | **0,486** | **0,439** | **0,475** |
| GRR RISK NEUT. | **0,246** | **0,311** | **0,486** | **1** | **0,989** | **0,986** |
| GRR RISK AV. | **0,179** | **0,264** | **0,439** | **0,989** | **1** | **0,968** |
| GRR RISK LOV. | **0,279** | **0,339** | **0,475** | **0,986** | **0,968** | **1** |

Source: own elaboration

***Table 21. Rank correlations across selected performances measures - measures based on quantiles. 36 monthly returns window***

| **36 RETURNS** | VR | VARR | STARR | GRR RISK NEUT. | GRR RISK AV. | GRR RISK LOV. |
|---|---|---|---|---|---|---|
| VR | **1** | **0,561** | **0,936** | **0,332** | **0,279** | **0,332** |
| VARR | **0,561** | **1** | **0,579** | **0,646** | **0,614** | **0,646** |
| STARR | **0,936** | **0,579** | **1** | **0,507** | **0,454** | **0,507** |
| GRR RISK NEUT. | **0,332** | **0,646** | **0,507** | **1** | **0,993** | **1,000** |
| GRR RISK AV. | **0,279** | **0,614** | **0,454** | **0,993** | **1** | **0,993** |
| GRR RISK LOV. | **0,332** | **0,646** | **0,507** | **1,000** | **0,993** | **1** |

Source: own elaboration

### 3.2.3.5 Measures derived from utility functions

The final group is referred to the **measures derived from utility functions**.

The rank correlations between **MRAR 2** and the other measures are, for all the windows, under the critical value (as claimed also by Caporin & Lisi (2011)).

The **MRAR 20** and **MRAR 50** result highly correlated only when the window is composed by 120 returns. However, the values of their correlations are always bigger than 0,832; in studies with a larger number of analysed assets it may be considered sufficient to allow for correlation. Bearing in mind this “bias” and in order to avoid the selection of three indicators of the same type we select the MRAR 2 and just one of the other two indicators.

***Table 22. Rank correlations across selected performances measures – measures derived from utility functions. 178 monthly returns window***

| **178 RETURNS** | MRAR 2 | MRAR 20 | MRAR 50 |
|---|---|---|---|
| MRAR 2 | **1** | **0,136** | **-0,046** |
| MRAR 20 | **0,136** | **1** | **0,893** |
| MRAR 50 | **-0,046** | **0,893** | **1** |

Source: own elaboration

***Table 23. Rank correlations across selected performances measures – measures derived from utility functions. 120 monthly returns window***

| 120 RETURNS | MRAR 2 | MRAR 20 | MRAR 50 |
|---|---|---|---|
| MRAR 2 | **1** | **0,625** | **0,504** |
| MRAR 20 | **0,625** | **1** | **0,932** |
| MRAR 50 | **0,504** | **0,932** | **1** |

Source: own elaboration

***Table 24. Rank correlations across selected performances measures – measures derived from utility functions. 60 monthly returns window***

| 60 RETURNS | MRAR 2 | MRAR 20 | MRAR 50 |
|---|---|---|---|
| MRAR 2 | **1** | **0,632** | **0,207** |
| MRAR 20 | **0,632** | **1** | **0,832** |
| MRAR 50 | **0,207** | **0,832** | **1** |

Source: own elaboration

***Table 25. Rank correlations across selected performances measures – measures derived from utility functions. 36 monthly returns window***

| 36 RETURNS | MRAR 2 | MRAR 20 | MRAR 50 |
|---|---|---|---|
| MRAR 2 | **1** | **0,529** | **0,214** |
| MRAR 20 | **0,529** | **1** | **0,857** |
| MRAR 50 | **0,214** | **0,857** | **1** |

Source: own elaboration

### 3.2.4 Rolling analysis of correlation

In order to confirm or invalidate some uncertain correlation detected with the static study, we present now a correlation analysis conducted following a **rolling approach**.

A correlation between two measures that results above the critical value in a static analysis may present sub-periods characterized by values below the threshold, and vice versa.

The Spearman rank correlations are now calculating applying **rolling windows of 60 months**, obtaining 118 snapshots of the rank correlation matrixes.

We will present and analyse just the abovementioned controversial cases.

#### 3.2.4.1 Interpretation of results

We now present the graphs relative to the rolling correlation between Treynor ratio and Sharpe ratio in **Figure 10** and between Treynor ratio and Jensen's alpha in **Figure 11**.

***Figure 10. Rolling rank correlation between Treynor ratio and Sharpe ratio***


Source: own elaboration

***Figure 11. Rolling rank correlation between Treynor ratio and Jensen's alpha***


Source: own elaboration

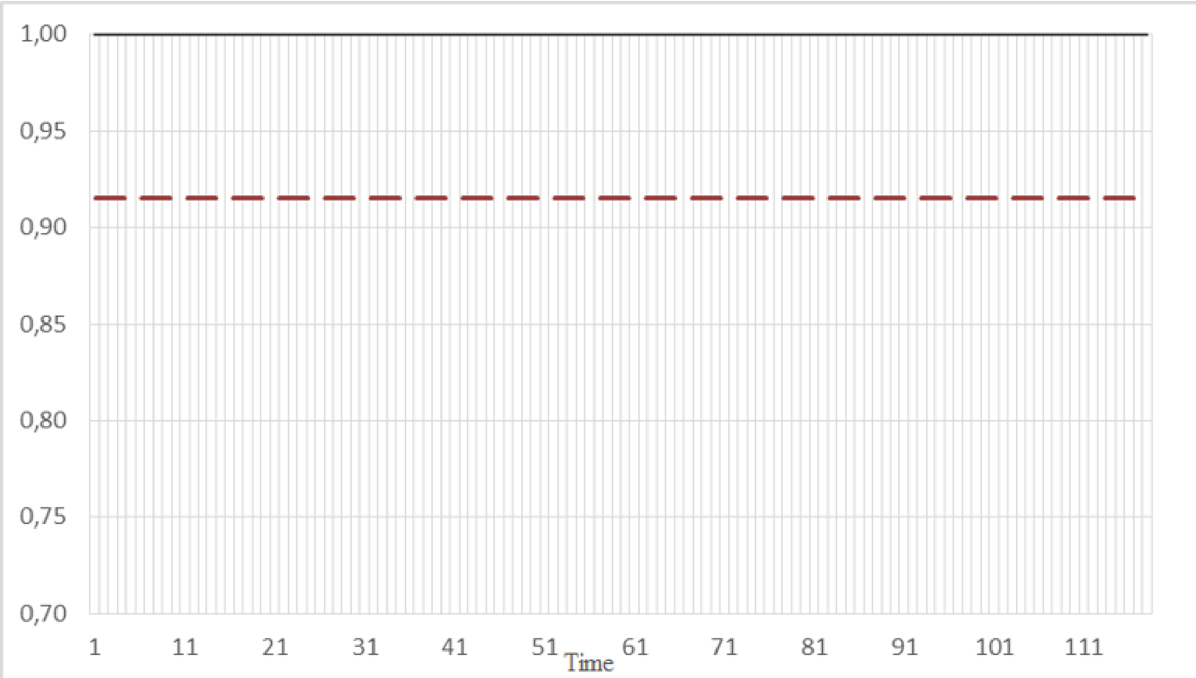
Regarding the correlation between **Treynor ratio** and **Sharpe ratio**, even though for some windows it goes under the threshold, usually it is very high and close to one. Moreover, it never goes below 0,814.

Concerning the second figure, **Treynor ratio** and **Jensen's alpha** provide a ranking correlation that goes under the critical value just in 14 out of 118 windows.

On the basis of these analyses and taking into consideration the bias of a very high threshold level due to a low number of examined assets, we exclude the **Treynor ratio** from the study. It does not incorporate different information from those carried by Sharpe ratio (as already claimed by Caporin & Lisi, (2011)) and Jensen's alpha.

The Figure 12 shows the correlation between **Sharpe ratio** and **$M^2$ measure**. As expected, it is always equal to one **(Figure 12)**.

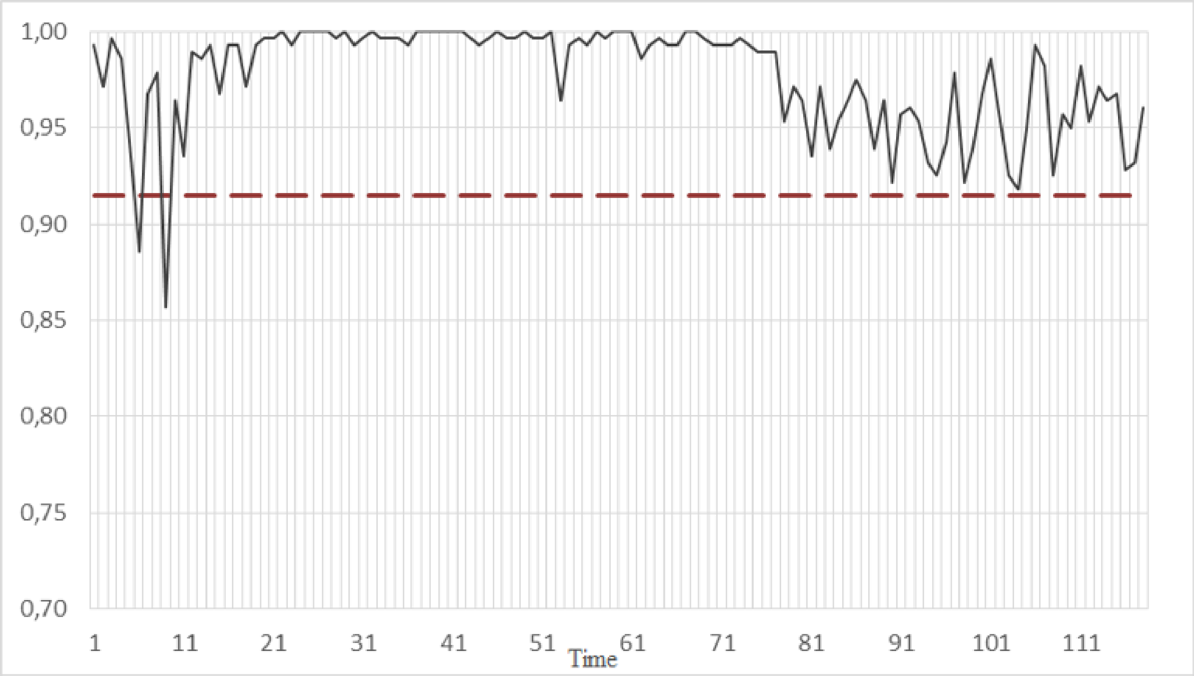
***Figure 12. Rolling rank correlation between Sharpe and M2 measure***




Source: own elaboration

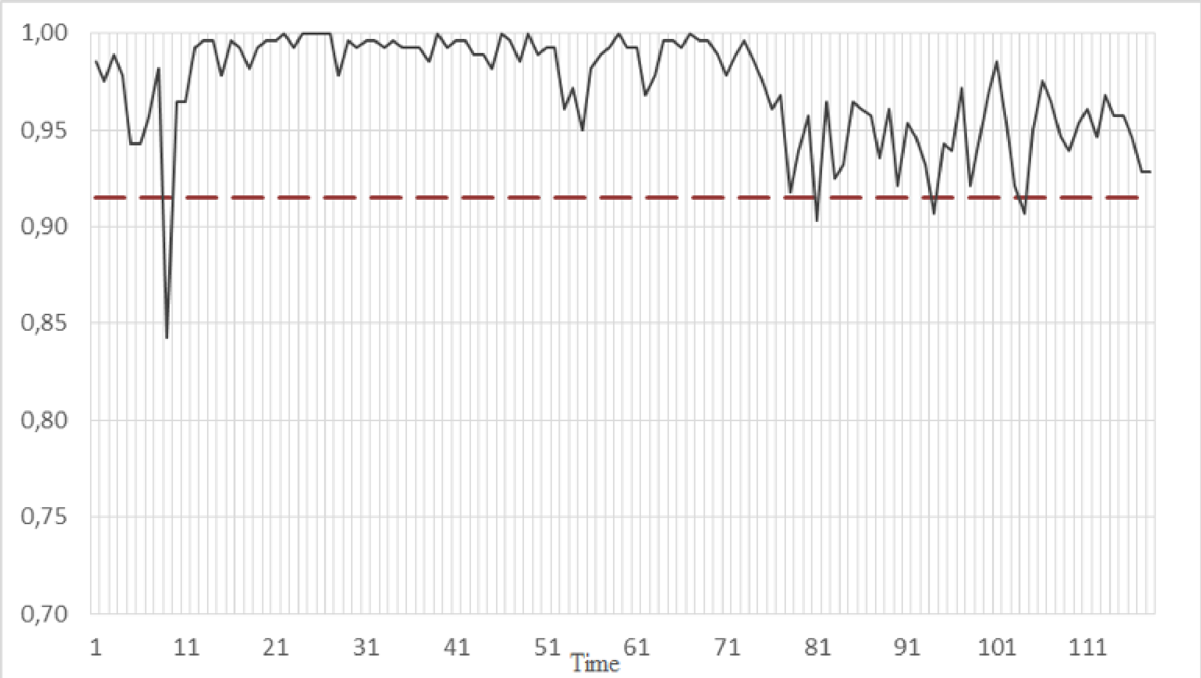
The following graphs **(Figure 13, Figure 14)** represent the rolling correlation between **Calmar ratio** and, respectively, **Sterling ratio** and **Martin ratio**.

***Figure 13. Rolling rank correlation between Calmar ratio and Sterling ratio***




Source: own elaboration

***Figure 14. Rolling rank correlation between Calmar ratio and Martin ratio***

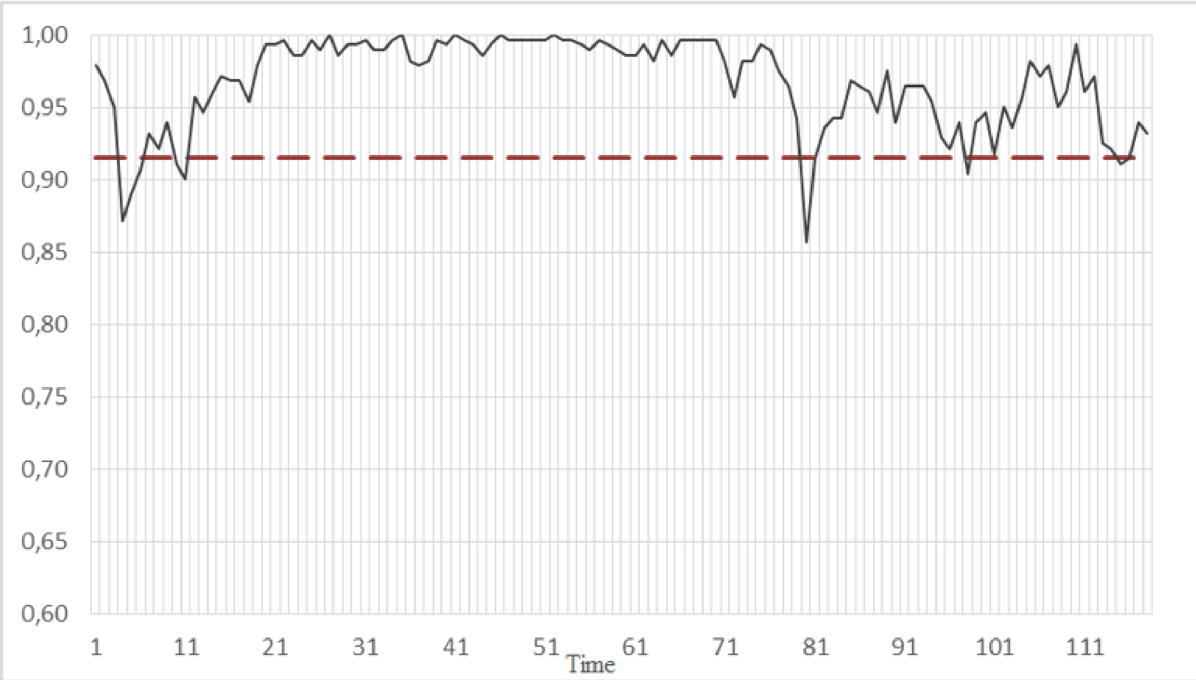



Source: own elaboration

From these figures we can conclude that, except for very few cases nonetheless characterized by high values (never lower than 0,842), the **Calmar ratio** provide ranking highly correlated with both **Sterling ratio** and **Martin ratio**. It appears reasonable to choose just **one indicator** belonging to this sub-group.

We now present the rolling correlation between the **Omega ratio** and the **K3 ratio (Figure 15)**. It shows a great stability above the threshold level (only 10 times lower than 0,915). For this reason, we will select only one between these two indicators.

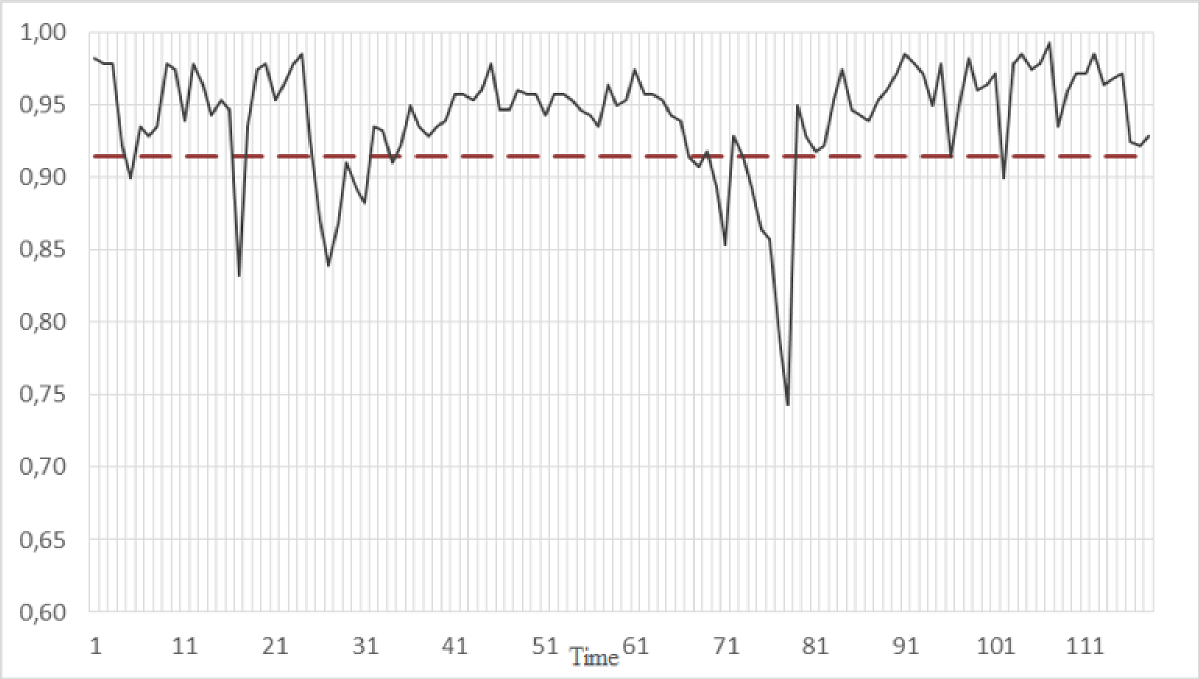
***Figure 15. Rolling rank correlation between Omega ratio and K3 ratio***




Source: own elaboration

In the next figure **(Figure 16),** we present the correlation between the **Farinelli- Tibiletti for a conservative investor** and the **Farinelli- Tibiletti for a "growth" investor**. Except for the spike occurred during the 78th window, the correlation values are always higher than 0,832. Bearing in mind the "bias" of a very high threshold level, we decide to introduce only one of these two indicator.

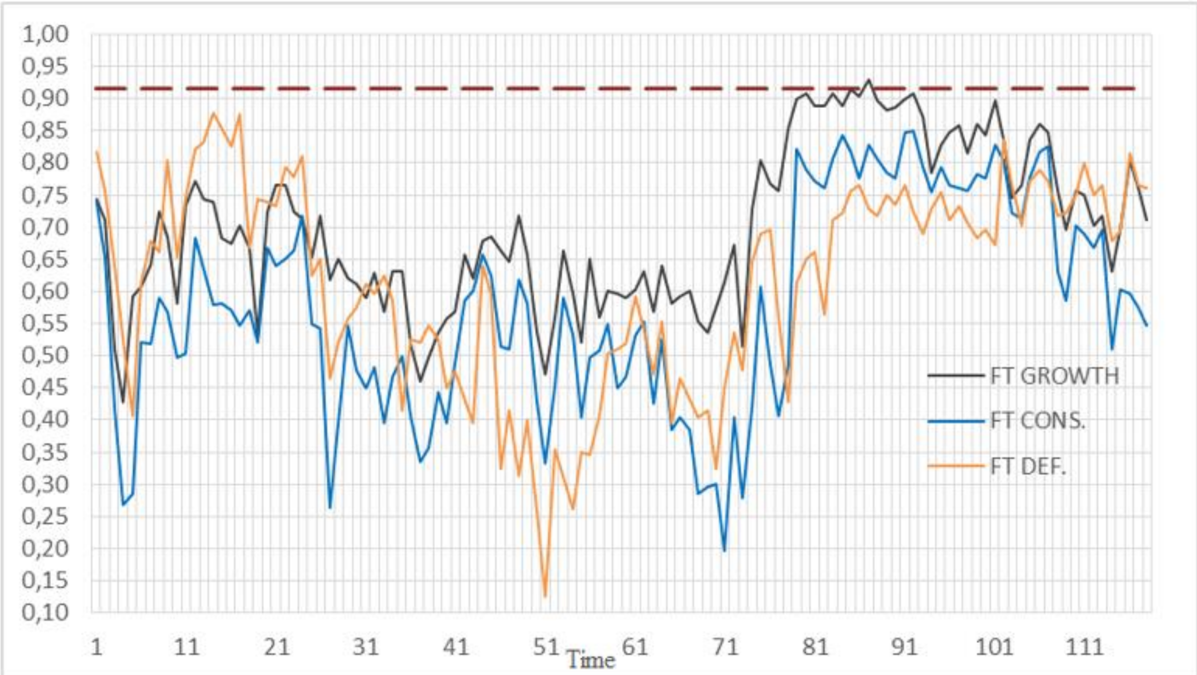
*Figure 16. Rolling rank correlation between FT for a conservative investor and for a "growth" investor*




Source: own elaboration

In order to confirm that the **Farinelli-Tibiletti for an aggressive investor** is not highly correlated with the **FT for other kinds of investors**, in **Figure 17** we exhibit their rolling correlations. They are clearly below the critical value (except for one window). This result is in line with the findings of Caporin & Lisi (2011).

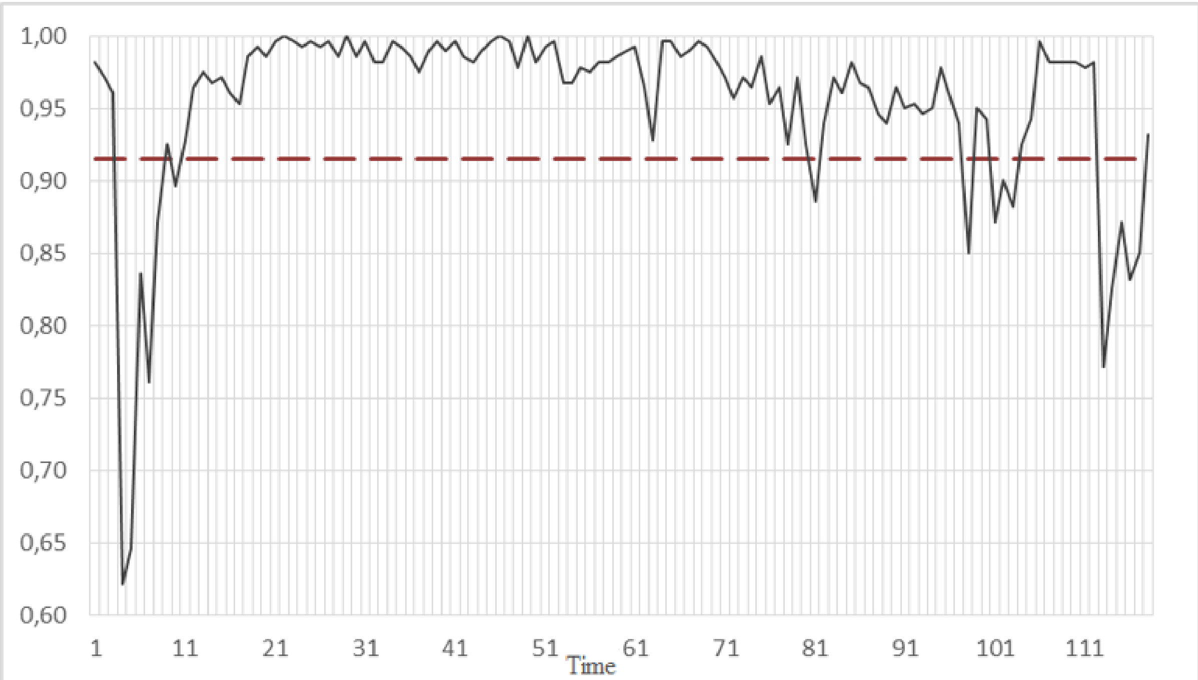
*Figure 17. Rolling rank correlation between FT for an aggressive investor and FT for other kind of investors*




Source: own elaboration

In **Figure 18,** we present the rolling correlation between the **VR ratio** and the **STARR ratio**. The majority of the correlation values are above the threshold; however, the static analysis probably released uncertain results for the spike at the beginning of the analysis. In order to provide a selection of measures which is limited, we decide to consider these two measures concordant.

***Figure 18. Rolling rank correlation between the VR ratio and STARR ratio***




Source: own elaboration

### 3.2.5 Conclusions about correlation feature

We now summarize the results of the **correlation analysis**.

Regarding the **traditional and other unclassified performance measures**, we will consider for the rest of the analysis the information ratio, the adjusted Sharpe ratio, the Diaman ratio and the Jensens's alpha. The M2 measure is excluded because perfectly correlated with the Sharpe ratio (and evidently with the other measures correlated with the Sharpe ratio). We will choose only one measure between the Sharpe ratio and the Treynor ratio.

Among the **measures based on drawdown**, characterized by a high within group correlation, we will select only one index representing the entire group.

Regarding the **measures based on partial moments**, we will consider only one indicator among the Sortino ratio, the K3 ratio, the Omega ratio and the Farinelli-Tibiletti for a defensive investor. We will take into consideration the FT for an aggressive investor for the rest of the analysis, as well as one measure between the FT for a conservative investor and the FT for a "growth" investor.

Regarding the **measures based on quantiles**, we will consider for the rest of the analysis the VARR ratio, one index between the VR ratio and the STARR ratio and one measure among the three GRR ratios.

Finally, regarding the **measures derived from utility functions**, we will take into account the MRAR 2 and one indicator between the MRAR 20 and the MRAR 50.

## 3.3 Stability of performance measures

### 3.3.1 Introduction

In order to select the final indicators, we calculate the **stability** of all the performance measures. At the end of this analysis, we will present, for each category of performance measures, a list of **uncorrelated** and **stable** indicators. Ultimately, these measures will be used as a linear combination to create a **composite indicator**, which maximizes the stability over time.

We combine two different approaches used by Menardi & Lisi: firstly, we will analyse the stability of performance measures calculated across different time horizons through the same stability index used in *"On the stability of performance measures over time: an empirical study"* (Menardi & Lisi, 2012a). Subsequently, we will try to increase the degree of stability of the single performance measures by constructing a composite indicator as in *"Are performance measures equally stable?"* (Menardi & Lisi, 2012b).

An important characteristic of a performance measure may be identified in its **stability** over time.

The persistence of the relative performance of a financial asset is the result of two components. The first component is the true persistence of the **financial asset performance**, which is the aptitude of the asset to repeat a positive behaviour over time. Taking a fund as an example, it derives from the positive results of the performances of the single securities in which the fund invests and, more in general, from the reference financial market. However, if we measure the performance using an unstable indicator, this feature may be partially hidden.

The second component is, therefore, the **stability of the performance measure** used to evaluate the performance of the financial asset. In this analysis, we will study this last aspect.

The choice of a specific performance indicator may be crucial for predictive purposes or, similarly, to anticipate the relative future behaviour of financial assets. The key point is that a performance measure characterized by a high stability may produce better evaluations of the true persistence in a relative sense (Bodson, Coën & Hübner, 2008).

### 3.3.2 Methodology used to analyse the stability

There are several ways to define the stability of a performance measure over time.

In this research, we use an index based on the **changes of funds ranking over time**. This index points out the degree of similarity among rankings induced by the different performance measures over time. The maximum stability occurs when a performance measures generates unchanging rankings over time.

In order to define the stability index, we need to set the following **variables**.

The funds are defined as $\mathbf{A_i}$, with i = [1,n] representing the entire set of analysed financial assets. Their returns are observed at different times, from 1 to T.

The entire sample T is subdivided into P contiguous sub-periods of length l, so that **T = P*l**. For example, a one year monthly time series (T=12) may be divided into two adjacent windows (P=2), each having a length of six months (l=6).

The different performance measures are indicated as M; $\mathbf{m_i^{(p)}}$ is the estimate of M over period p for the asset $A_i$, with p = [1,P].

$\mathbf{R_i^{(p)}}$ is the rank of asset $A_i$, among the entire sample of assets $A_{1,\dots,n}$, induced by the performance measure M in the period p.

The difference between rankings of a financial asset $A_i$ induced by a performance measure M, from period (p-1) to period p, is equal to $\mathbf{d_i^{(p)}(m)} = R_i^{(p)} - R_i^{(p-1)}$.

According to the described notation, we define the stability index of a performance measure M as follows:

$$I_0(M) = \frac{1}{P-1}\sum_{p=2}^{P}\frac{\sum_{i=1}^{n}\left|d_i^{(p)}\right|^q}{\psi}$$

, with $\psi$ as a normalizing factor.

If q is set to one, the numerator becomes a linear function; alternatively, if q is set to two, it turns into a quadratic form. The quadratic function has the feature of weighting large variations more than small variations; however, the risk is that few very large variations finish dominating the index. For this reason, we choose to use the linear function, setting **q equal to one**.

Therefore, the numerator is the sum of the absolute values of the differences between the rankings of n financial assets, induced by the analysed performance measure M, from p=2 to p=P.

We set the **normalizing factor** $\Psi$ equal to the maximum value of $\sum_i \left| d_i^{(p)} \right|$, so that $\Psi = \frac{1}{2} n^2$.

When ranks totally reverse from one period to the successive, $I_0(M)$ results equal to 1. On the contrary, when the ranks do not change over time, the stability index is equal to 0; this means that $0 \leq I_0(M) \leq 1$.

Summarizing, $I_0(M)$ is the average of the normalized sum of the absolute values of the differences between ranks of different assets over different periods.

In order to ensure that a high stability correspond to a high value of the index, we introduce **I(M) = 1 - $I_0$(M).**

I(M) takes its **minimum value (zero)** when the rankings are inverted over adjacent periods **(unstable indicator)** and its **maximum value (one)** when the rankings do not change over time **(stable indicator)**.

### 3.3.3 Application on performance measures

We now calculate the stability indices for each performance measure using the 15 selected funds.

Their monthly returns are divided into 2, 4 and 8 adjacent time windows (P = {2, 4, 8}), respectively of length of 88, 44 and 22 returns (l = {88, 44, 22}). In order to take into account the same set of data for all the time windows, we exclude the first two returns of the entire sample of 178 returns. In this way, T is equal to 176.

The **Table 26** reports the stability indices of all the performance measure, for the three different choices of l and P.

*Table 26. Stability index for performance measures and for different l and P*

| | P=2, L=88 | P=4, L=44 | P=8, L=22 | AVERAGE |
|---|---|---|---|---|
| SHARPE | 0,3244 | 0,3304 | 0,4159 | 0,3569 |
| TREYNOR | 0,3778 | 0,2830 | 0,4006 | 0,3538 |
| ALPHA J | 0,3600 | 0,3659 | 0,3905 | 0,3721 |
| INF. RATIO | 0,6800 | 0,3541 | 0,4235 | 0,4859 |
| M2 | 0,3244 | 0,3304 | 0,4159 | 0,3569 |
| ADJ. SHARPE | 0,2533 | 0,1822 | 0,3473 | 0,2610 |
| DIAMAN | 0,5022 | 0,3481 | 0,3448 | 0,3984 |
| CALMAR | 0,2356 | 0,2830 | 0,3981 | 0,3055 |
| STERLING | 0,3778 | 0,2889 | 0,4159 | 0,3608 |
| BURKE | 0,3778 | 0,2948 | 0,3956 | 0,3560 |
| MARTIN | 0,4133 | 0,3185 | 0,3930 | 0,3750 |
| SORTINO | 0,3244 | 0,3422 | 0,4057 | 0,3575 |
| K3 | 0,3600 | 0,3244 | 0,4006 | 0,3617 |
| OMEGA | 0,3244 | 0,3481 | 0,4108 | 0,3611 |
| FT DEF. | 0,3600 | 0,3007 | 0,4032 | 0,3546 |
| FT CONS. | 0,5378 | 0,2119 | 0,4057 | 0,3851 |
| FT GROWTH | 0,3778 | 0,2711 | 0,4133 | 0,3541 |
| FT AGGR. | 0,2356 | 0,3778 | 0,4057 | 0,3397 |
| VR | 0,3600 | 0,3244 | 0,4133 | 0,3659 |
| VARR | 0,1822 | 0,2059 | 0,4006 | 0,2629 |
| STARR | 0,3600 | 0,3126 | 0,3879 | 0,3535 |
| GRR RISK NEUT. | 0,4133 | 0,2237 | 0,3422 | 0,3264 |
| GRR RISK AV. | 0,4133 | 0,2356 | 0,3422 | 0,3304 |
| GRR RISK LOV. | 0,3067 | 0,2178 | 0,3422 | 0,2889 |
| MRAR 2 | 0,5200 | 0,3481 | 0,3448 | 0,4043 |
| MRAR 20 | 0,5378 | 0,4963 | 0,4260 | 0,4867 |
| MRAR 50 | 0,6622 | 0,5200 | 0,5352 | 0,5725 |

Source: own elaboration

### 3.3.3.1 Interpretation of general results

The above table clearly shows that the stability of each indicator changes depending on the **choice of P and l**. This is in line with the findings of Menardi & Lisi (2012a,b). On average, the case in which the sample is divided into 8 sub-periods (P=8) results the one with the most stable indices. The choice of the window length should depend on the time horizon of the investor. However, the study suggest that, on average, the performance measures express the maximum stability over a 22 months period. This choice of width l seems to provide a high stability to the indicators: in Menardi & Lisi (2012b) the performance measures reach the maximum stability when l=24, whereas in Menardi & Lisi (2012a) this happens with l=21 and l=15, in line with our result of l=22.

Analysing the different **categories of performance measures**, the group derived from utility functions (MRAR) appear one of the most stable, confirming what was found in Menardi & Lisi (2012a,b).

In particular, the **indicator** for extremely risk adverse investors MRAR 50 is, on average, the most stable measure (in Menardi & Lisi (2012a,b) the most stable indicator is the Appraisal ratio, not analysed in our study). On the contrary, the Sharpe ratio adjusted for skewness and kurtosis turn out to be the less stable.

We now underline an interesting aspect emerged from the analysis. Regarding the measures for which some specific parameter has to be set depending on the **risk appetite of the investor**, those built for risk averse investors are, on average, more stable. This fact happens to the Farinelli-Tibiletti ratios, whose conservative and defensive versions are on average the most stable formats (Menardi & Lisi (2012a) consistently stated that the FT for aggressive investors is the most unstable measure of the FT ratios). Regarding the Generalized Rachev Ratios, once again, the risk averse version is the most stable one. Finally, regarding the MRAR ratios, increasing the risk aversion, increasing the stability.

#### 3.3.3.2 Choice of final performance measures

In order to select for each group of performance measures the restricted group indicators, we now select the **most stable measures, on average, among those** previously found **correlated**. Between the Sharpe ratio and the Treynor ratio, we thus choose the Sharpe ratio.

Among the measures based on drawdown, we select the Martin ratio.

Regarding the measures based on partial moments, among the Sortino ratio, the K3 ratio, the Omega ratio and the Farinelli-Tibiletti for a defensive investor, we pick the K3 ratio. Between the FT for a conservative investor and the FT for a “growth” investor, we select the first index. Between the VR ratio and the STARR ratio, we prefer the VR ratio, whereas among the three different GRR ratios we select the version for risk averse investors.

Between the MRAR 20 and the MRAR 50, we choose the MRAR 50.

Finally, we exclude from the analysis the adjusted for skewness and kurtosis Sharpe ratio because it appears as the most unstable measure.

Summarizing, we now present the group of performance measures selected to create the composite indicator: the Sharpe ratio, the Jensen's alpha, the information ratio, the Diaman ratio, the Martin ratio, the K3 ratio, the FT ratio for a conservative investor, the FT ratio for an aggressive investor, the VR ratio, the VARR ratio, the GRR for a risk averse investor, the MRAR 2 and the MRAR 50.

### 3.3.4 Composite indicator in order to maximize stability

In order to measure the performances of a financial asset, the selection of a single indicator may be a limited approach. For this reason, the combination of different measures would be desirable in order to obtain a **composite indicator** that summarizes the information included in the single measures. The aim is to construct an index that is **as stable as possible**.

#### 3.3.4.1 Theory aspects

The composite indicator is built as a linear combination of single performance measures and it is defined as M*=a'M, where $a=(a_1,\dots,a_d)$ is the weight vector. Considering the constraints, the problem may be set down as follows:

$$M^* = \hat{a}'M = \sum_{j=1}^{d} \hat{a}_j M_j \text{ with } \begin{cases} \hat{a} = \operatorname{argmax}_a I(a'M) \\ a'a = 1 \\ a \geq 0 \end{cases} .$$

Therefore, the composite indicator M* is constructed as the linear combination of "d" single performance measures and the weight vector "a" is the one that maximizes the stability of the composite indicator. There are two rational constraints beneath the optimization problem: the sum of the weights assigned to single performance measures has to be equal to one (a'a=1) and all the weights have to be positive (a≥0).

In order to transform all the performance measures with the same size scale, for each sub-period all the indicators have to be standardized in a number from 0 to 1 with a **cross-section normalization**. First, we calculate the performance measures for the 15 funds in every sub-period. Then, every performance measure is normalized by subtracting to the indicator the minimum value of the measure in that sub-period and dividing this dividend by the difference between the maximum and the minimum values of the indicator in that sub-period. For example, to be clearer, in a given sub-period the Sharpe ratio is calculated for all the funds. In order to transform Sharpe ratios into numbers from 0 to 1, the smaller Sharpe ratio value calculated in that sub-period is subtracted to the single Sharpe ratio value, and the resulting difference is divided by the maximum Sharpe ratio minus the minimum Sharpe ratio calculated in that period. Recalling that $m_i^{(p)}$ is the estimate of performance measure M over period p for asset $A_i$, the normalized value of $m_i^{(p)}$ is equal to $m^*{}_i^{(p)} = \frac{m_i^{(p)} - \text{MIN}\,(m_{i,\dots,n}^{(p)})}{\text{MAX}\,(m_{i,\dots,n}^{(p)}) - \text{MIN}\,(m_{i,\dots,n}^{(p)})}$ .

Then, in every sub-period, we calculates for each fund the composite indicator as the **weighted sum of the normalized performance measures.**

Finally, we constructs the fund **ranks** based on these composite indicators.

The final stability index I(M*) is calculated using these rankings:

$$I(M^*) = 1 - \frac{1}{P-1}\sum_{p=2}^{P}\frac{\sum_{i=1}^{n}\left|d_{M^*}{}^{(p)}\right|}{\frac{1}{2}n^2} .$$

In order to maximize the stability, we choose the weight composition (assigned to the single performance measures) which maximize the stability index I(M*).

### 3.3.4.2 Maximization results

In order to maximize the stability of the composite indicator, for each choice of l and P, we run **1.040.460 random simulations** of the possible **weight compositions**.

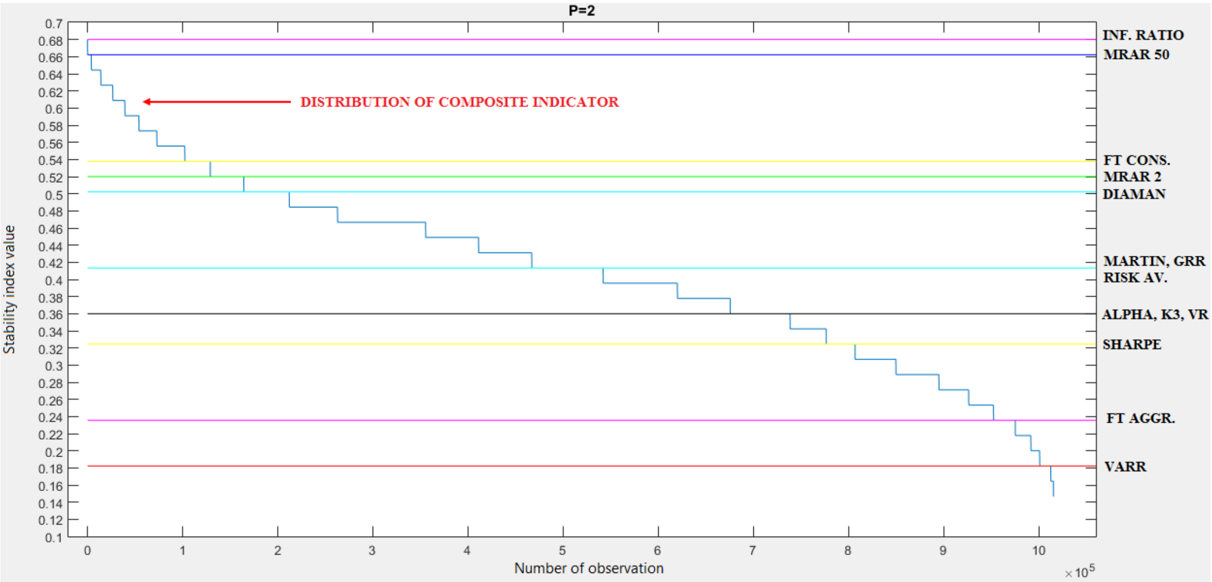
The composite index will have at least the same stability of most stable single performance measure. In fact, if a gain in stability compared to the single measures is not possible, there will be a scenario assigning the total weight to the most stable measure.

In the next figures, we present the **sorted distribution** of the stability of the composite indicator in the **1.040.460 scenarios**, for the different choices of P and l. In particular, the first figure **(Figure 19)** is referred to the case of P=2 and l=88, the second one **(Figure 20)** to the case of P=4 and l=44 and the third one **(Figure 21)** to the case of P=8 and l=22.

In order to facilitate a comparison between the stability of the composite indicator and the stability of the single measures, we point out the stability of each performance measure in all the figures.

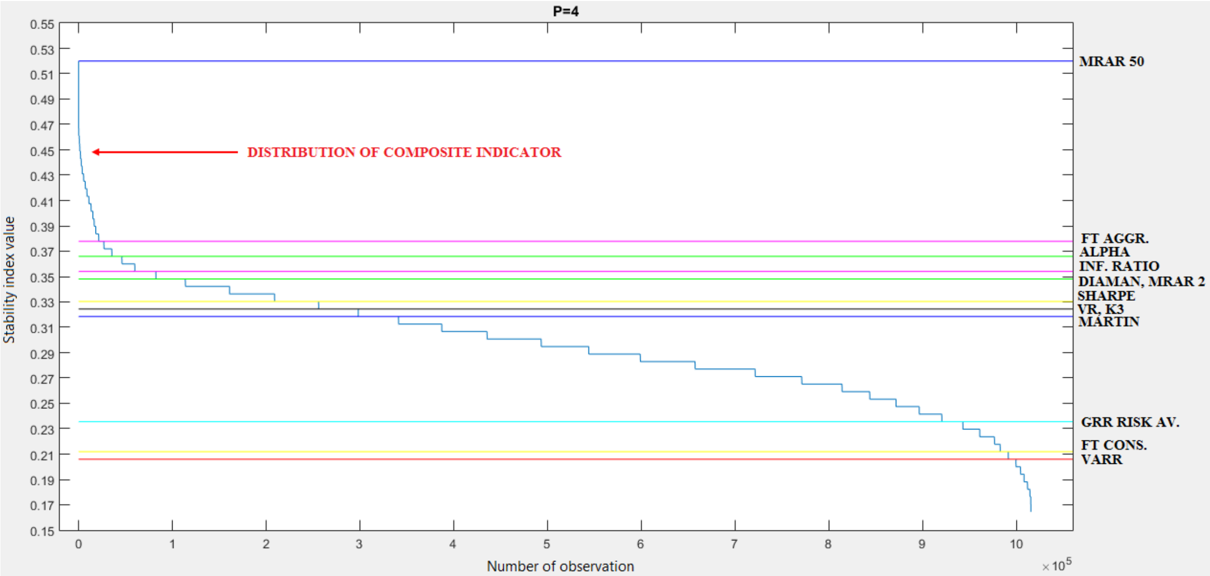
A first evidence shows that in two out of three cases the composite indicator is not able to have a gain in terms of stability compared to the single indicators, whereas it **improve the stability** of the most stable performance measure when P=8 and l=22 (Figure 21). Moreover, in all the cases, the lowest stability of the 1.040.460 composite indicators is always “worst” than the stability of the most unstable measure, meaning that a diversification is not necessarily good.

***Figure 19. Distribution of stability index for composite indicator and stability of single measures – P=2, l=88***



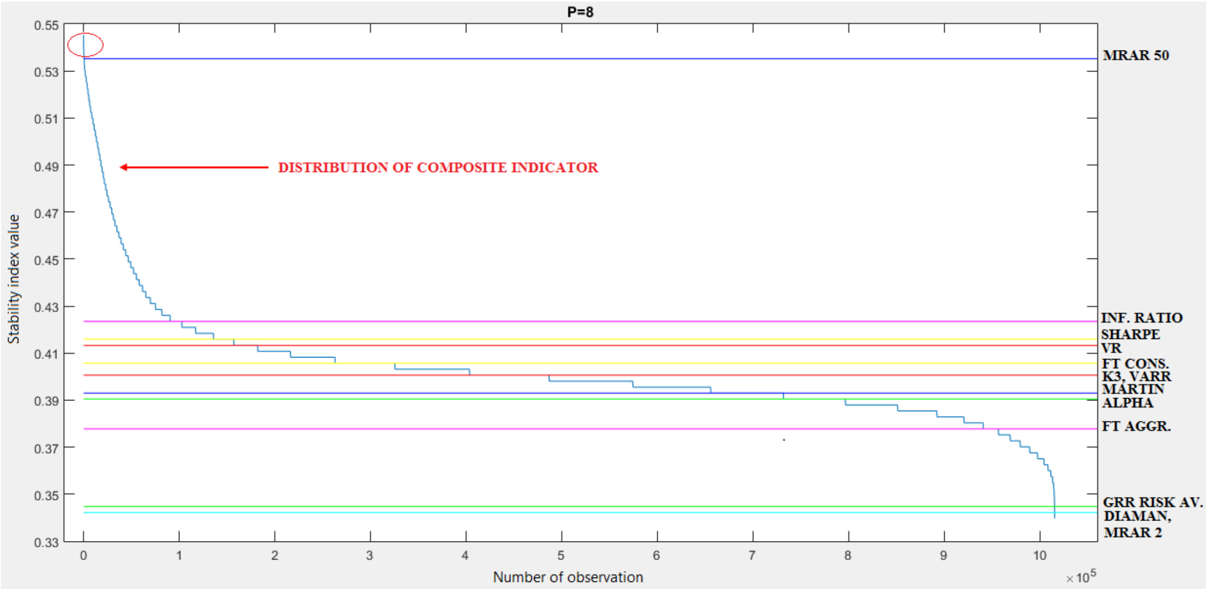

Source: own elaboration

***Figure 20. Distribution of stability index for composite indicator and stability of single measures – P=4, l=44***




Source: own elaboration

***Figure 21. Distribution of stability index for composite indicator and stability of single measures – P=8, l=22***




Source: own elaboration

Regarding the **first case (when P=2 and l=88)**, we can highlight that the most stable composite indicator has the same stability of the most stable indicator, the information ratio. However, we can show that different compositions brings to that value. The next table **(Table 27)** sets out **50 compositions** of the composite indicator with a stability index of **0,6800**, the maximum value. All these combinations assign to **MRAR 50** or to **information ratio** an important weight. Interestingly, the MRAR 50, which is not the most stable measure when P=2 and l=88, is improving its stability when combined with other measures, reaching the maximum value of stability of that period.

***Table 27. Compositions of composite indicator which maximize stability – P=2, l=88***

| P=2, L=88 | Variable Name | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Simulation # | 1_Sharpe | 2_AlphaJ | 3_Inf. Ratio | 4_Diaman | 5_Martin | 6_K3 | 7_FT. Conserv. | 8_FT. Aggress. | 9_Vr | 10_Varr | 11_GRR. Risk Aver. | 12_MRAR 2 | 13_MRAR 50 | STABILITY |
| 1 | 0% | 0% | 1% | 0% | 0% | 0% | 0% | 0% | 0% | 5% | 0% | 0% | 94% | **0,6800** |
| 2 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 7% | 0% | 0% | 93% | **0,6800** |
| 3 | 0% | 2% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 4% | 1% | 2% | 91% | **0,6800** |
| 4 | 0% | 0% | 0% | 0% | 0% | 0% | 4% | 0% | 0% | 6% | 0% | 0% | 90% | **0,6800** |
| 5 | 2% | 0% | 0% | 0% | 0% | 0% | 0% | 1% | 0% | 5% | 3% | 0% | 89% | **0,6800** |
| 6 | 6% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 4% | 1% | 1% | 88% | **0,6800** |
| 7 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 7% | 2% | 1% | 81% | **0,6800** |
| 8 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 7% | 3% | 1% | 80% | **0,6800** |
| 9 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 8% | 4% | 1% | 78% | **0,6800** |
| 10 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 8% | 5% | 1% | 77% | **0,6800** |
| 11 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 8% | 6% | 1% | 76% | **0,6800** |
| 12 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 3% | 11% | 3% | 74% | **0,6800** |
| 13 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 4% | 10% | 4% | 73% | **0,6800** |
| 14 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 2% | 14% | 1% | 73% | **0,6800** |
| 15 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 2% | 14% | 2% | 72% | **0,6800** |
| 16 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 17% | 2% | 71% | **0,6800** |
| 17 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 3% | 13% | 3% | 71% | **0,6800** |
| 18 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 17% | 3% | 70% | **0,6800** |
| 19 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 3% | 13% | 4% | 70% | **0,6800** |
| 20 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 1% | 16% | 3% | 70% | **0,6800** |
| 21 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 5% | 10% | 7% | 69% | **0,6800** |
| 22 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 17% | 4% | 69% | **0,6800** |
| 23 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 3% | 13% | 5% | 69% | **0,6800** |
| 24 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 1% | 16% | 4% | 69% | **0,6800** |
| 25 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 4% | 11% | 8% | 68% | **0,6800** |
| 26 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 3% | 14% | 6% | 68% | **0,6800** |
| 27 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 15% | 6% | 68% | **0,6800** |
| 28 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 16% | 6% | 68% | **0,6800** |
| 29 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 16% | 5% | 68% | **0,6800** |
| 30 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 17% | 5% | 68% | **0,6800** |
| 31 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 4% | 12% | 6% | 68% | **0,6800** |
| 32 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 1% | 16% | 5% | 68% | **0,6800** |
| 33 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 16% | 7% | 67% | **0,6800** |
| 34 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 17% | 6% | 67% | **0,6800** |
| 35 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 2% | 15% | 6% | 67% | **0,6800** |
| 36 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 1% | 16% | 6% | 67% | **0,6800** |
| 37 | 1% | 1% | 1% | 1% | 1% | 1% | 1% | 2% | 1% | 1% | 16% | 7% | 66% | **0,6800** |
| 38 | 0% | 0% | 100% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | **0,6800** |
| 39 | 0% | 0% | 99% | 0% | 0% | 0% | 0% | 1% | 0% | 0% | 0% | 0% | 0% | **0,6800** |
| 40 | 0% | 0% | 98% | 0% | 0% | 0% | 0% | 1% | 0% | 0% | 0% | 1% | 0% | **0,6800** |
| 41 | 0% | 0% | 84% | 0% | 0% | 0% | 0% | 4% | 0% | 11% | 0% | 0% | 1% | **0,6800** |
| 42 | 0% | 0% | 80% | 0% | 0% | 0% | 0% | 20% | 0% | 0% | 0% | 0% | 0% | **0,6800** |
| 43 | 11% | 0% | 79% | 0% | 10% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | **0,6800** |
| 44 | 0% | 0% | 79% | 0% | 16% | 0% | 0% | 5% | 0% | 0% | 0% | 0% | 0% | **0,6800** |
| 45 | 0% | 0% | 78% | 0% | 0% | 0% | 0% | 21% | 0% | 0% | 0% | 1% | 0% | **0,6800** |
| 46 | 0% | 0% | 75% | 10% | 0% | 0% | 0% | 14% | 0% | 0% | 0% | 0% | 1% | **0,6800** |
| 47 | 0% | 0% | 72% | 3% | 20% | 0% | 1% | 0% | 0% | 4% | 0% | 0% | 0% | **0,6800** |
| 48 | 4% | 0% | 70% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 26% | 0% | **0,6800** |
| 49 | 1% | 0% | 70% | 2% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 27% | 0% | **0,6800** |
| 50 | 0% | 0% | 69% | 1% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 30% | 0% | **0,6800** |

Source: own elaboration

With regard to the **second case (when P=4 and l=44)**, the most stable composite indicator has the same stability of the most stable indicator, the MRAR 50. Looking at the weight compositions, the maximum value of stability is reached in four occasions **(Table 28)**. In these cases, at least **97% of the weight** is assigned to the most stable measure, the MRAR 50: the final indicator is more similar to a **single performance measure** than to a composite index.

***Table 28. Compositions of composite indicator which maximize stability – P=4, l=44***

| P=4, L=44 | Variable Name | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Simulation # | 1_Sharpe | 2_AlphaJ | 3_Inf. Ratio | 4_Diaman | 5_Martin | 6_K3 | 7_FT. Conserv. | 8_FT. Aggress. | 9_Vr | 10_Varr | 11_GRR. Risk Aver. | 12_MRAR 2 | 13_MRAR 50 | STABILITY |
| 1 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 100% | **0,5200** |
| 2 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 1% | 0% | 0% | 0% | 0% | 99% | **0,5200** |
| 3 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 2% | 0% | 0% | 0% | 0% | 98% | **0,5200** |
| 4 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 3% | 0% | 0% | 0% | 0% | 97% | **0,5200** |

Source: own elaboration

Finally, when **P=8 and l=22** the composite indicator has a **gain in terms of stability** compared to the single measures. The composite indicators built with the weight combinations showed in **Table 29** reach a stability value of **0,5556**, whereas the most stable performance measure (MRAR 50) has a stability of **0,5352.**

The displayed compositions are not the only ones to increase the stability compared to the MRAR 50. However, we report them since they bring to the **highest value of stability**.

The first composition assign the 87% of the total weight to the MRAR 50, followed by the VARR ratio, the Farinelli Tibiletti ratio for an aggressive investor and the MRAR 2, weighting 7%, 4% and 2% respectively. The second composition differs only in the percentage assigned to the MRAR 50 and to the FT for an aggressive investor, weighting 86% and 5% respectively.

***Table 29. Compositions of composite indicator which maximize stability – P=8, l=22***

| P=8, L=22 | Variable Name | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Simulation # | 1_Sharpe | 2_AlphaJ | 3_Inf. Ratio | 4_Diaman | 5_Martin | 6_K3 | 7_FT. Conserv. | 8_FT. Aggress. | 9_Vr | 10_Varr | 11_GRR. Risk Aver. | 12_MRAR 2 | 13_MRAR 50 | STABILITY |
| 1 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 4% | 0% | 7% | 0% | 2% | 87% | **0,5556** |
| 2 | 0% | 0% | 0% | 0% | 0% | 0% | 0% | 5% | 0% | 7% | 0% | 2% | 86% | **0,5556** |

Source: own elaboration

Therefore, is there any benefit in terms of stability in the use of a **composite indicator**? Summarizing, we would answer **yes** just for the **last case**.

The study suggests that for **long-time investments** (88 months and 44 months, corresponding to P=2 and P=4), the use of a composite indicator does not bring to a gain in terms of stability compared to the single performance measures. On the contrary, for a **short-time investment** (22 months, corresponding to P=8), the construction of a mixed indicator leads to an improvement in terms of stability. In this case, there is no single measure that is more stable than the composite index.

In the **next chapter**, both single measures and composite indicators will be tested as instruments for the selection of mutual funds. In order to be able to analyse the composite indicators, for each choice of P and l we have to select only a **single composition** among those bringing to the maximum stability, even though they are theoretically equivalent. The preferred **weight compositions** are summarized in the next table **(Table 30).**

*Table 30. Compositions of composite indicator which maximize stability*

| | P=2, L=88 | P=4, L=44 | P=8, L=22 |
|---|---|---|---|
| SHARPE | 1% | 0% | 0% |
| ALPHA J | 1% | 0% | 0% |
| INF. RATIO | 1% | 0% | 0% |
| DIAMAN | 1% | 0% | 0% |
| MARTIN | 1% | 0% | 0% |
| K3 | 1% | 0% | 0% |
| FT CONS. | 1% | 0% | 0% |
| FT AGGR. | 2% | 3% | 5% |
| VR | 1% | 0% | 0% |
| VARR | 1% | 0% | 7% |
| GRR RISK AV. | 16% | 0% | 0% |
| MRAR 2 | 6% | 0% | 2% |
| MRAR 50 | 67% | 97% | 86% |
| STABILITY INDEX | 0,6800 | 0,5200 | 0,5556 |

Source: own elaboration

## 3.4 Conclusions

In the **first part** of this chapter, we developed an analysis of the possible **within group mutual correlations** between the performance measures. The correlations were examined using the Spearman rank correlation. First, we carried out a **static analysis** realized over four different evaluation windows. Then, in order to confirm or invalidate some uncertain correlation, we applied a **rolling approach**.

In the **second part** of the chapter, we developed an analysis of the **stability** of the performance measure based on changes of fund rankings over time. In order to select the performance measures among those correlated, we calculated the **stabilities** of all the **single indicators** across different time horizons. **Among the correlated measures**, we chose the **most stable** ones.

Through the **13 selected performance measures** (the Sharpe ratio, the Jensen's alpha, the information ratio, the Diaman ratio, the Martin ratio, the K3 ratio, the FT ratio for a conservative investor, the FT ratio for an aggressive investor, the VR ratio, the VARR ratio, the GRR for risk averse investors, the MRAR 2 and the MRAR 50), we realized an attempt to increase the degree of stability of the single indicators by constructing a **composite index**. It was only when P=8 and l=22 that the composite indicator had a **gain in term of stability** compared to the single measures.

In the next chapter, these composite indicators and the single performance measures will be finally analysed through a fund selection process.

# CHAPTER FOUR

# 4 Implementation of the composite indicator and of the single performance measures for a fund selection process

## 4.1 Introduction

In the previous chapters, we implemented both a descripted analysis of the performance measures and a study of some of their features, such as correlation and stability. We now present a **fund selection process** based on these indices.

In particular, we test all the single performance measures and the three different composite indicators (one for each time window) in order to verify which measure lead to a better fund selection. Moreover, we add an equally weighted portfolio to this comparison as a model for a naïve approach.

The aim is to build up funds of funds portfolios with rolling evaluations of every performance measure and then to analyse certain characteristics of these portfolios.

## 4.2 Set of financial instruments used for the analysis

In order to verify some result obtained in the prior chapter, the following study is conducted on **different funds** compared to those previously used.

In particular, we select **generic** balanced funds euro denominated; they may be composed by a higher equity component or by a higher fixed-income component (theoretically they can invest from 10% to 90% of the portfolio in stocks). Once again, we choose maximum one fund for each investment company.

In order to obtain from the analysis results as robust as possible, we use only funds with at least 10 year of history. Specifically, the first date of common history is the 17/02/2004 and the final date is the 01/09/2017. For each fund, we thus obtain 163 monthly prices, equivalent to 162 monthly return (from 31/03/2004 to 31/08/2017).

The benchmark, as in the previous analysis, is a composite index 50% composed of the Bloomberg Barclays Euro Aggregate Bond Index and 50% of the FTSE All World Index. The risk-free rate is set to zero.

The **final 13 funds** are displayed in the following table (**Table 31**), including the benchmark.

***Table 31. Second selection of funds and benchmark***

| N° | ISIN | Fund Denomination | Currency | Fund Inception Date |
|---|---|---|---|---|
| 1 | IT0003081525 | ALLIANZ MULTIPARTNER C.TO MULTI50 | EUR | 02/05/2001 |
| 2 | FR0000294308 | AMUNDI PORTFOLIO STRATEGIE OBLIG 5-7 EURO | EUR | 20/01/2004 |
| 3 | IT0000380706 | ANIMA VISCONTEO A | EUR | 20/05/1985 |
| 4 | IT0000380003 | ARCA BB | EUR | 18/09/1984 |
| 5 | LU0180870494 | AZ FUND 1 C.TO CONSERVATIVE | EUR | 17/02/2004 |
| 6 | LU0049912065 | CANDRIAM BIL PATRIMONIAL C.TO HIGH | EUR | 15/02/1994 |
| 7 | FR0010149211 | CARMIGNAC GESTION FCP PROFIL REACTIF 100 A ACC. | EUR | 02/01/2002 |
| 8 | LU0134132231 | EUROFUND LUX C.TO IPAC BALANCED | EUR | 23/11/2001 |
| 9 | LU0140420323 | FTIF TEMPLETON GLOBAL BALANCED N ACC. | EUR | 31/12/2001 |
| 10 | LU0115099839 | JPM INVEST.FUND C.TO GLOBAL BALANCED | EUR | 10/07/2002 |
| 11 | LU0090850685 | LEMANIK ITALY A CAP. | EUR | 30/12/1998 |
| 12 | IT0004764491 | SYMPHONIA PATRIMONIO REDDITO R | EUR | 03/11/1999 |
| 13 | LU0167296127 | UBS (LUX) STRATEGY FUND C.TO GROWTH (EUR) | EUR | 26/06/2003 |
| 14 | BENCHMARK | 50% BARCLAYS EUROAGG TR INDEX - 50% FTSE ALL WORLD INDEX | | |

Source: own elaboration

## 4.3 Portfolio series construction: methodology description

We now present the **methodology** used to build up the portfolios based on the single performance measures and on the composite indicators.

The entire sample of returns T is subdivided into P adjacent time windows of length l. In particular, 160 monthly returns are divided into 2, 4 and 8 contiguous sub-periods (P = {2, 4, 8}) respectively of length 80, 40 and 20 returns (l = {80, 40, 20}). We thus exclude from the analysis the first two data of the entire sample of 162 monthly returns.

After every sub-period, we calculate all the performance measures on the funds, based on the last “l” returns. Following each evaluation, for every performance measures we select the **best three funds**. We assign a weight of 50% to the best fund, a weight of 30% to the second best fund and a weight of 20% to the third best fund.

For every performance measure, the **return series of the fund of funds portfolio** is developed summing up the weighted returns of the best three funds, from the successive returns compared to when the evaluation is computed until the new evaluation (after “l” returns).

For each of the three different choices of P, we finally obtain: 1) “n” portfolio series built up with the best funds according to the “n” single performance measures; 2) a portfolio series built up with the best funds according to the composite indicator; 3) an equally weighted portfolio.

***Example of the methodology***: portfolio series based on Sharpe ratio when P=8 and l=20.

After the first 20 returns (l=20), the Sharpe ratios are calculated on all the funds. The best three funds are selected. The **“Sharpe” portfolio series** starts from the 21$^{st}$ return and it is built up as a weighted sum of the best three funds (50% of the weight to the best fund, 30% to the second best fund and 20% to the third best fund). This weights hold until the new evaluation based on the Sharpe ratio. After 40 returns from the beginning, the Sharpe ratios are calculated on the previous 20 returns. The best three funds are selected and, from the successive return, the "Sharpe" portfolio series is built up with the new weights. This procedure continues until the end of the sample.

The result is a portfolio series built up with the best funds based on the Sharpe ratio.

## 4.4 Comparison of the performances of the portfolio series

In the rest of this section, we present an analysis of the results emerged from the construction of the portfolios through the performance measures, the composite indicators and the equally weighted portfolio.

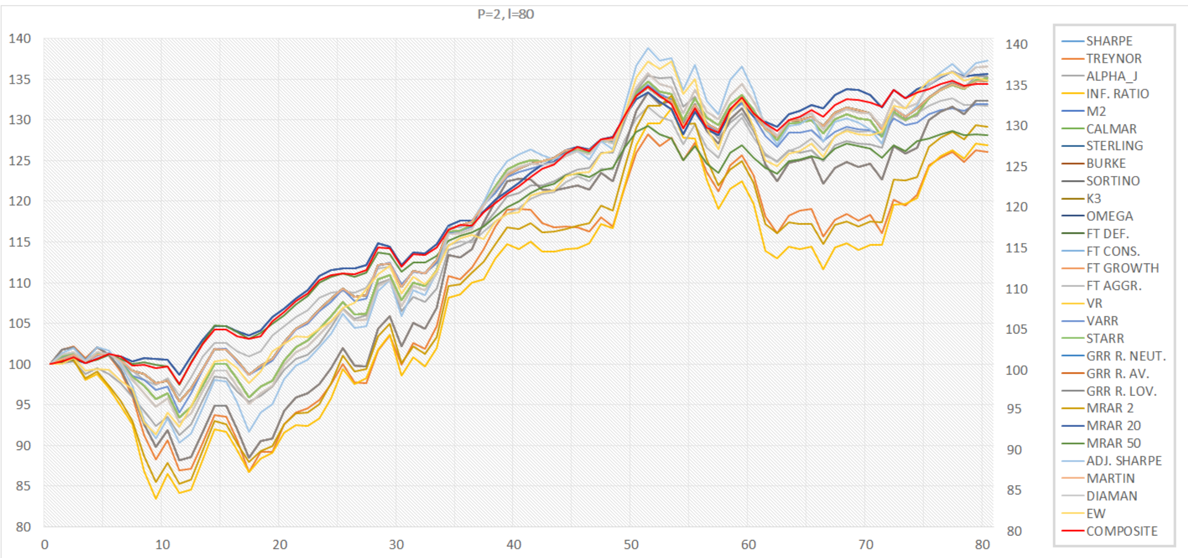
At first, we compare the cumulated returns of the funds of funds portfolios and their turnover rates.

Subsequently, other portfolios characteristics in terms of risk and return are analysed (average annualized return, annualized volatility, maximum drawdown and CVaR 95%). We create a ranking of the portfolios for each of these features and a final composite ranking.

### 4.4.1 Analysis of cumulated returns and turnover rates

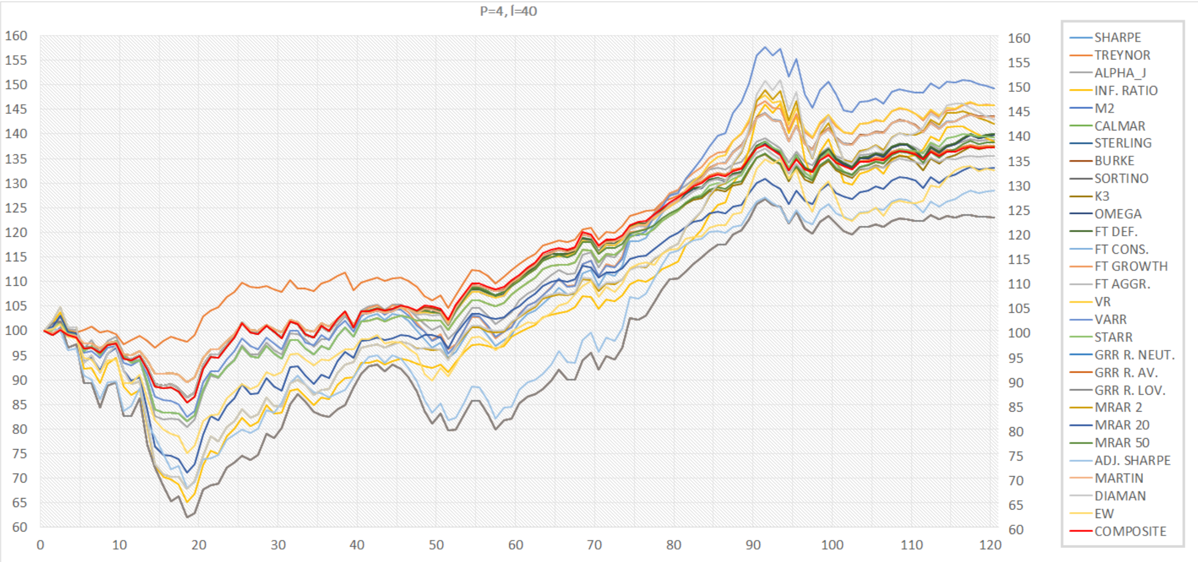
**Figure 22**, **Figure 23** and **Figure 24** show the **cumulative returns** of the portfolios built through the different measures, respectively for P=2, P=4 and P=8.

*Figure 22. Time series of cumulative returns – P=2, l=80*



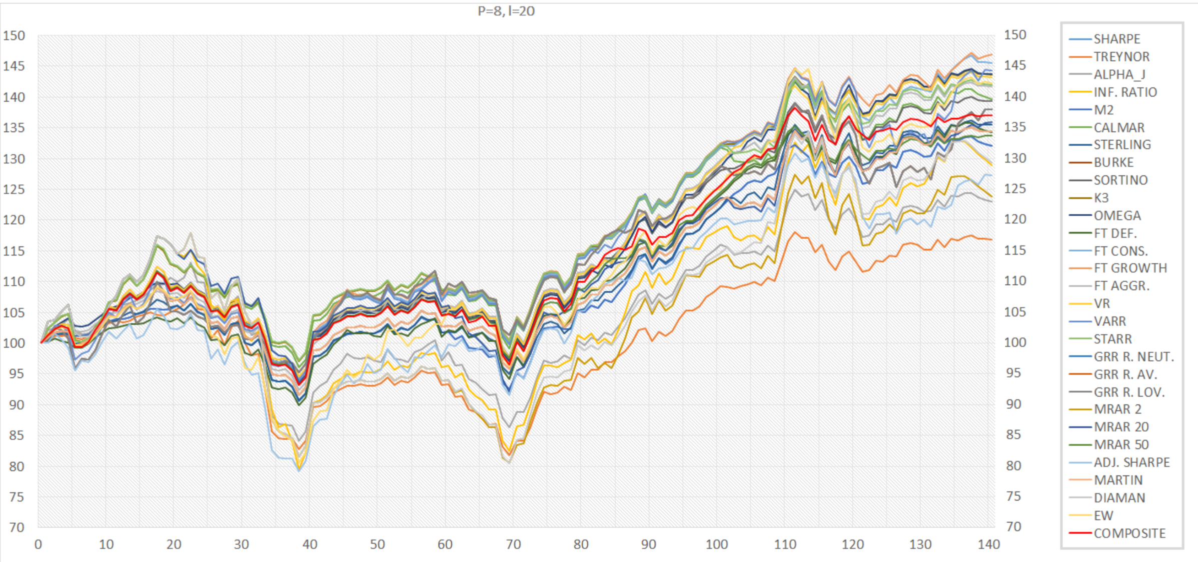
Source: own elaboration

***Figure 23. Time series of cumulative returns – P=4, l=40***




Source: own elaboration

***Figure 24. Time series of cumulative returns – P=8, l=20***




Source: own elaboration

In the next table (**Table 32**) we present the **final cumulated returns** and the **turnover rates** for every portfolio series.

In particular, the **turnover rate** is calculated as follows: no matter the position within the best three funds, if from one sub-period to the successive an instrument, before excluded, enters in the portfolio composed by the best three funds, it counts as a change. If two new funds enter, we consider two changes. If all the funds are new compared to the portfolio of the previous sub-period, we consider three changes. Vice versa, if the three funds selected by a performance measure remain the same, no changes are counted.

On the basis of these arguments, the **maximum** possible changes in the composition of the portfolio in terms of fund is equal to the maximum possible changes from one sub-period to the successive (three funds), multiplied by the total amount of sub-periods minus two. In formula, 3*(l-2). We exclude two sub-periods firstly because the weights deriving by the evaluation of a performance measure on the last window are not applied to construct any fund of funds; secondly, the first composition is by definition new (without changes).

Thus, the turnover rate is calculated as the ratio between the sum of the changes on the best three funds from one period to the other and the maximum amount of all possible changes. An indicator that in every sub-period select different funds as best three has a turnover rate of 100%, whereas an indicator that select always the same three best funds has a turnover rate of 0%. Thus, lower the turnover rate, higher the stability of the performance measure in selecting the best three funds.

Note that when P=2 and l=80 all the turnover rates are equal to 0. In fact, the first evaluation by a performance measure is realized on the first 80 returns and the resulting weights are applied for the successive 80 returns (from the $81^{st}$ to the $160^{th}$), until the end of the sample. Thus, by construction, there are no changes in the best three funds.

*Table 32. Cumulated returns and turnover rate of portfolio series*

| | P=2, l=80 | | P=4, l=40 | | P=8, l=20 | |
|---|---|---|---|---|---|---|
| | CUMULATED RETURN | TURNOVER RATE | CUMULATED RETURN | TURNOVER RATE | CUMULATED RETURN | TURNOVER RATE |
| SHARPE | 34,87% | 0,0% | 39,96% | 33,3% | 32,03% | 66,7% |
| TREYNOR | 26,05% | 0,0% | 37,61% | 50,0% | 16,84% | 72,2% |
| ALPHA J | 35,36% | 0,0% | 38,76% | 50,0% | 22,95% | 72,2% |
| INF. RATIO | 26,87% | 0,0% | 38,36% | 83,3% | 28,92% | 66,7% |
| M2 | 34,87% | 0,0% | 39,96% | 33,3% | 32,03% | 66,7% |
| ADJ. SHARPE | 37,26% | 0,0% | 28,51% | 66,7% | 27,22% | 77,8% |
| DIAMAN | 36,59% | 0,0% | 42,85% | 83,3% | 29,38% | 72,2% |
| CALMAR | 34,87% | 0,0% | 39,96% | 33,3% | 39,66% | 94,4% |
| STERLING | 35,12% | 0,0% | 43,44% | 50,0% | 35,45% | 77,8% |
| BURKE | 34,87% | 0,0% | 43,58% | 50,0% | 34,42% | 83,3% |
| MARTIN | 34,87% | 0,0% | 43,44% | 50,0% | 34,42% | 83,3% |
| SORTINO | 34,87% | 0,0% | 39,77% | 33,3% | 39,28% | 83,3% |
| K3 | 35,12% | 0,0% | 37,47% | 50,0% | 43,60% | 83,3% |
| OMEGA | 35,65% | 0,0% | 39,77% | 33,3% | 43,71% | 66,7% |
| FT DEF. | 34,87% | 0,0% | 39,96% | 33,3% | 34,27% | 66,7% |
| FT CONS. | 35,12% | 0,0% | 43,49% | 66,7% | 45,49% | 83,3% |
| FT GROWTH | 36,59% | 0,0% | 45,89% | 50,0% | 46,90% | 72,2% |
| FT AGGR. | 31,71% | 0,0% | 35,45% | 50,0% | 41,72% | 77,8% |
| VR | 35,12% | 0,0% | 45,82% | 33,3% | 43,19% | 72,2% |
| VARR | 31,93% | 0,0% | 49,30% | 50,0% | 44,33% | 72,2% |
| STARR | 35,12% | 0,0% | 39,30% | 50,0% | 42,01% | 83,3% |
| GRR RISK NEUT. | 32,41% | 0,0% | 23,07% | 83,3% | 37,94% | 77,8% |
| GRR RISK AV. | 32,41% | 0,0% | 23,07% | 83,3% | 37,94% | 77,8% |
| GRR RISK LOV. | 32,41% | 0,0% | 23,07% | 83,3% | 37,94% | 77,8% |
| MRAR 2 | 29,15% | 0,0% | 42,12% | 83,3% | 23,87% | 72,2% |
| MRAR 20 | 35,65% | 0,0% | 33,03% | 66,7% | 35,92% | 83,3% |
| MRAR 50 | 28,09% | 0,0% | 38,29% | 50,0% | 33,66% | 61,1% |
| EW | 34,87% | 0,0% | 32,58% | 0,0% | 42,16% | 0,0% |
| COMPOSITE | 34,45% | 0,0% | 37,29% | 33,3% | 37,01% | 72,2% |

Source: own elaboration

The **turnover rate**, calculated as previously described, may be interpreted as a sort of **stability**; it measures the stability of a performance measure in selecting the best three funds. For this reason, we carry out a comparison with the results about stability obtained in the previous chapter.

The fact that the **stability** of each indicator change depending on the **choice of l and P** is confirmed in the analysis of the turnover rate. However, the case in which the sample is divided into 8 sub-periods results with a larger turnover rate compared to the other case (whereas, in the previous analysis, the maximum stability was reached when P was equal to 8).

The **MRAR 50** is confirmed to be, on average, one of the most stable measures.

Moreover, the aspect underlined in the previous analysis that the indicators built for **risk averse investors** are, on average, more stable, is here confirmed. Both regarding the MRAR indicators and the Farinelli-Tibiletti ratios, the versions built for risk averse investors result those with the lowest turnover rates.

With P=4 the **composite indicator** proved to be the indicator with the lowest turnover rate. Thus, it was able improve the stability of the portfolios based on the single performance measures.

Analysing the **turnover rates** and the **cumulated returns** jointly, we point out an important aspect: a low turnover rate (high stability) does not always imply high returns.

In terms of **cumulated returns**, the VR ratio, the VARR ratio and the Farinelli-Tibiletti ratios for growth investors are among the indicators with the highest cumulated returns, for all the possible choices of P and l.

### 4.4.2 Analysis of risk and return statistics

In the following tables (**Table 33**, **Table 34** and **Table 35**), we present some **statistics** referred to all the analysed portfolios.

In particular, the first column expresses the **average annualized return** of the portfolios. It is calculated as $[(1+\mu)^{12} - 1]$, where μ is the mean of the portfolio returns.

The second column is the **annualized volatility**. It is calculated as $[\sigma * \sqrt{12}]$, where σ is the standard deviation of the portfolio series.

The third column shows the **maximum drawdown**, whereas the forth exhibits the **CVaR** at **95% confidence level.**

We developed a ranking among the portfolios for each of the statistics described above. Clearly, for the statistics about risk it holds the rule that smaller the value, better the ranking. The final ranking is built as an average of the other ranking, re-ranked.

Looking at the results, it is evident that all the portfolios are characterized by **positive annualized average returns**, for all the choices of P and l. However, even the average returns of the equally weighted portfolios are always positive.

From a comparison of all the indicators, it appears that the best measure **in terms of returns** is the **Farinelli-Tibiletti for growth investors**. Whatever the choice of P and l, it is among the measures with a better ranking in terms of returns.

Regarding the **volatility**, among the category of the traditional performance measures the Sharpe ratio and the M2 measure are, on average, the best indicators. The measures based on drawdown are characterized by good rankings in terms of volatility. On average, these measures are those with the best rankings in terms of volatility among all the categories of indicators. Among all, Burke ratio and Martin ratio positively stand out.

As regards to the measures based on partial moments, the Omega ratio and the FT for defensive investors result in having the best volatility rankings. Finally, concerning the measures derived from utility function, the MRAR 50 have on average the best ranking. It is interesting to note that among the abovementioned measures, many indicators are built for risk averse investors (i.e. FT for defensive investors and MRAR 50). These results show that these measures are characterized, on average, by lower volatility compared to the measures built for risk lover investors.

Analysing the statistics on **maximum drawdown** and **CVaR at 95% confidence level**, the rankings seem to be very similar to those generated by the annualized volatility. The best indicators for each category of performance measures are, on average, the same abovementioned for volatility. In particular: the Sharpe ratio and the M2 measure for the traditional measures, the Burke ratio and the Martin ratio for the measures based on drawdown, the Omega ratio and the FT for defensive investors for the measures based on partial moment and the MRAR 50 for the measures derived from utility function. The results seem to certify that these measures tend to select financial instruments characterized by relative low risk: as a consequence, they have better ranks both in terms of volatility, maximum drawdown and CVaR at 95% confidence level.

This analysis highlights an **interesting aspect**. The best performance measures in terms of returns are often among the worst in terms of risks (volatility, maximum drawdown and CVaR 95%), and vice versa. The Farinelli-Tibiletti for growth investors is a good example: when P=2, P=4 and P=8 it is respectively the second best, the third best and the best portfolio in terms of returns. However, in terms of volatility, it is respectively the twentieth, the seventeenth and the fifteenth best. These results seem to certify that, in order to obtain higher returns, an indicator has to invest in riskier assets.

In the analysis of the **final ranking**, we have to take into account that it is built as the average of the previous rankings. Considering that three out of the four statistics are referred to the risk attitude of the performance measure (volatility, maximum drawdown and CVaR 95%), the final ranking promotes the indices which tend to select "low risk" funds. Thus, the best measures in terms of final ranking are the same previously described as the best in terms of volatility, maximum drawdown and CVaR 95% (among all, Sharpe ratio, Burke ratio, Martin ratio, Omega ratio and FT for defensive investors).

The portfolios based on the **composite indicators**, as it was previously analysed, presents a good behaviour in terms of turnover rates. Regarding the statistics on risk and return, when P=2 the portfolio based on the composite indicator is among the best in terms of volatility, maximum drawdown and CVaR 95%, but among the worst in terms of returns. On the other hand, the other two composite indicators (when P=4 and P=8) express average behaviours, both regarding returns and risk performances. The portfolios of the composite indicators, therefore, do not show a unique conduct in terms of risk and return.

***Table 33. Statistics and rankings of portfolio series – P=2, l=80***

| P=2, l=80 | ANNUALIZED RETURN | ANNUALIZED VOLATILITY | MAX DRAWDOWN | CVAR 95% | RANKING RETURN | RANKING VOLATILITY | RANKING DRAWDOWN | RANKING CVAR 95% | FINAL RANKING |
|---|---|---|---|---|---|---|---|---|---|
| **SHARPE** | 4,69% | 4,43% | -5,82% | -2,32% | 13 | 6 | 6 | 6 | **3** |
| **TREYNOR** | 3,85% | 7,86% | -15,14% | -4,38% | 28 | 28 | 28 | 28 | **28** |
| **ALPHA J** | 4,79% | 5,35% | -9,02% | -2,77% | 4 | 19 | 21 | 21 | **21** |
| **INF. RATIO** | 3,98% | 8,19% | -16,97% | -4,94% | 27 | 29 | 29 | 29 | **29** |
| **M2** | 4,69% | 4,43% | -5,82% | -2,32% | 13 | 6 | 6 | 6 | **3** |
| **ADJ. SHARPE** | 5,14% | 7,33% | -11,55% | -3,84% | 1 | 23 | 23 | 26 | **23** |
| **DIAMAN** | 4,94% | 5,45% | -8,53% | -2,64% | 2 | 20 | 19 | 14 | **18** |
| **CALMAR** | 4,69% | 4,43% | -5,82% | -2,32% | 13 | 6 | 6 | 6 | **3** |
| **STERLING** | 4,76% | 5,19% | -7,88% | -2,65% | 6 | 14 | 14 | 16 | **13** |
| **BURKE** | 4,69% | 4,43% | -5,82% | -2,32% | 13 | 6 | 6 | 6 | **3** |
| **MARTIN** | 4,69% | 4,43% | -5,82% | -2,32% | 13 | 6 | 6 | 6 | **3** |
| **SORTINO** | 4,69% | 4,43% | -5,82% | -2,32% | 13 | 6 | 6 | 6 | **3** |
| **K3** | 4,76% | 5,19% | -7,88% | -2,65% | 6 | 14 | 14 | 16 | **13** |
| **OMEGA** | 4,74% | 3,48% | -3,94% | -1,95% | 11 | 2 | 1 | 1 | **1** |
| **FT DEF.** | 4,69% | 4,43% | -5,82% | -2,32% | 13 | 6 | 6 | 6 | **3** |
| **FT CONS.** | 4,76% | 5,19% | -7,88% | -2,65% | 6 | 14 | 14 | 16 | **13** |
| **FT GROWTH** | 4,94% | 5,45% | -8,53% | -2,64% | 2 | 20 | 19 | 14 | **18** |
| **FT AGGR.** | 4,30% | 4,06% | -5,19% | -2,18% | 25 | 5 | 5 | 5 | **12** |
| **VR** | 4,76% | 5,19% | -7,88% | -2,65% | 6 | 14 | 14 | 16 | **13** |
| **VARR** | 4,35% | 4,60% | -7,21% | -2,42% | 24 | 13 | 13 | 13 | **20** |
| **STARR** | 4,76% | 5,19% | -7,88% | -2,65% | 6 | 14 | 14 | 16 | **13** |
| **GRR RISK NEUT.** | 4,58% | 7,42% | -13,63% | -3,80% | 21 | 24 | 24 | 23 | **24** |
| **GRR RISK AV.** | 4,58% | 7,42% | -13,63% | -3,80% | 21 | 24 | 24 | 23 | **24** |
| **GRR RISK LOV.** | 4,58% | 7,42% | -13,63% | -3,80% | 21 | 24 | 24 | 23 | **24** |
| **MRAR 2** | 4,21% | 7,62% | -15,09% | -4,22% | 26 | 27 | 27 | 27 | **27** |
| **MRAR 20** | 4,74% | 3,48% | -3,94% | -1,95% | 11 | 2 | 1 | 1 | **1** |
| **MRAR 50** | 3,84% | 3,43% | -4,51% | -2,00% | 29 | 1 | 4 | 3 | **11** |
| **EW** | 4,77% | 5,86% | -9,37% | -3,70% | 5 | 22 | 22 | 22 | **22** |
| **COMPOSITE** | 4,61% | 3,66% | -4,21% | -2,05% | 20 | 4 | 3 | 4 | **3** |

Source: own elaboration

***Table 34. Statistics and rankings of portfolio series – P=4, l=40***

| P=4, l=40 | ANNUALIZED RETURN | ANNUALIZED VOLATILITY | MAX DRAWDOWN | CVAR 95% | RANKING RETURN | RANKING VOLATILITY | RANKING DRAWDOWN | RANKING CVAR 95% | FINAL RANKING |
|---|---|---|---|---|---|---|---|---|---|
| **SHARPE** | 3,55% | 5,08% | -13,77% | -2,85% | 11 | 7 | 6 | 7 | **6** |
| **TREYNOR** | 3,32% | 3,84% | -6,44% | -2,26% | 22 | 1 | 1 | 1 | **5** |
| **ALPHA J** | 3,54% | 6,43% | -21,68% | -4,16% | 16 | 18 | 20 | 20 | **20** |
| **INF. RATIO** | 3,73% | 9,14% | -37,16% | -7,18% | 10 | 25 | 26 | 26 | **23** |
| **M2** | 3,55% | 5,08% | -13,77% | -2,85% | 11 | 7 | 6 | 7 | **6** |
| **ADJ. SHARPE** | 2,95% | 8,92% | -32,28% | -5,94% | 26 | 23 | 23 | 23 | **26** |
| **DIAMAN** | 4,07% | 9,17% | -35,25% | -7,07% | 2 | 26 | 24 | 25 | **22** |
| **CALMAR** | 3,55% | 5,08% | -13,77% | -2,85% | 11 | 7 | 6 | 7 | **6** |
| **STERLING** | 3,78% | 4,50% | -10,69% | -2,60% | 8 | 3 | 2 | 3 | **2** |
| **BURKE** | 3,78% | 4,45% | -10,69% | -2,57% | 7 | 2 | 2 | 2 | **1** |
| **MARTIN** | 3,78% | 4,50% | -10,69% | -2,60% | 8 | 3 | 2 | 3 | **2** |
| **SORTINO** | 3,54% | 5,09% | -13,77% | -2,85% | 17 | 11 | 6 | 7 | **10** |
| **K3** | 3,40% | 5,77% | -19,89% | -3,69% | 20 | 15 | 17 | 16 | **19** |
| **OMEGA** | 3,54% | 5,09% | -13,77% | -2,85% | 17 | 11 | 6 | 7 | **10** |
| **FT DEF.** | 3,55% | 5,08% | -13,77% | -2,85% | 11 | 7 | 6 | 7 | **6** |
| **FT CONS.** | 3,90% | 6,55% | -19,89% | -3,91% | 6 | 19 | 17 | 19 | **15** |
| **FT GROWTH** | 4,03% | 5,95% | -13,77% | -3,11% | 3 | 17 | 6 | 15 | **10** |
| **FT AGGR.** | 3,21% | 4,97% | -13,77% | -2,79% | 23 | 6 | 6 | 6 | **10** |
| **VR** | 3,95% | 4,57% | -10,69% | -2,69% | 5 | 5 | 2 | 5 | **4** |
| **VARR** | 4,33% | 6,81% | -18,01% | -3,87% | 1 | 20 | 16 | 18 | **14** |
| **STARR** | 3,55% | 5,85% | -19,89% | -3,76% | 15 | 16 | 17 | 17 | **18** |
| **GRR RISK NEUT.** | 2,63% | 10,16% | -39,13% | -7,47% | 27 | 27 | 27 | 27 | **27** |
| **GRR RISK AV.** | 2,63% | 10,16% | -39,13% | -7,47% | 27 | 27 | 27 | 27 | **27** |
| **GRR RISK LOV.** | 2,63% | 10,16% | -39,13% | -7,47% | 27 | 27 | 27 | 27 | **27** |
| **MRAR 2** | 4,00% | 8,95% | -35,25% | -6,94% | 4 | 24 | 24 | 24 | **21** |
| **MRAR 20** | 3,19% | 7,56% | -30,91% | -5,59% | 24 | 22 | 22 | 22 | **25** |
| **MRAR 50** | 3,43% | 5,22% | -14,77% | -2,96% | 19 | 14 | 14 | 14 | **15** |
| **EW** | 3,11% | 7,02% | -25,83% | -4,74% | 25 | 21 | 21 | 21 | **24** |
| **COMPOSITE** | 3,35% | 5,13% | -14,77% | -2,94% | 21 | 13 | 14 | 13 | **15** |

Source: own elaboration

***Table 35. Statistics and rankings of portfolio series – P=8, l=20***

| P=8, l=20 | ANNUALIZED RETURN | ANNUALIZED VOLATILITY | MAX DRAWDOWN | CVAR 95% | RANKING RETURN | RANKING VOLATILITY | RANKING DRAWDOWN | RANKING CVAR 95% | FINAL RANKING |
|---|---|---|---|---|---|---|---|---|---|
| **SHARPE** | 2,52% | 4,72% | -14,77% | -3,03% | 24 | 1 | 5 | 1 | **3** |
| **TREYNOR** | 1,51% | 5,77% | -24,57% | -4,38% | 29 | 21 | 23 | 23 | **24** |
| **ALPHA J** | 2,00% | 6,40% | -25,54% | -4,70% | 28 | 24 | 25 | 25 | **25** |
| **INF. RATIO** | 2,54% | 8,15% | -32,24% | -5,49% | 23 | 29 | 29 | 27 | **28** |
| **M2** | 2,52% | 4,72% | -14,77% | -3,03% | 24 | 1 | 5 | 1 | **3** |
| **ADJ. SHARPE** | 2,34% | 7,02% | -24,97% | -4,96% | 26 | 26 | 24 | 26 | **25** |
| **DIAMAN** | 2,57% | 8,08% | -31,79% | -5,95% | 22 | 28 | 28 | 29 | **27** |
| **CALMAR** | 3,07% | 5,71% | -17,08% | -3,46% | 10 | 17 | 21 | 17 | **21** |
| **STERLING** | 2,76% | 5,06% | -15,31% | -3,19% | 17 | 8 | 9 | 10 | **10** |
| **BURKE** | 2,70% | 5,08% | -16,01% | -3,09% | 18 | 9 | 15 | 5 | **12** |
| **MARTIN** | 2,70% | 5,08% | -16,01% | -3,09% | 18 | 9 | 15 | 5 | **12** |
| **SORTINO** | 3,01% | 5,03% | -15,54% | -3,08% | 11 | 5 | 10 | 4 | **2** |
| **K3** | 3,29% | 5,21% | -16,15% | -3,23% | 4 | 13 | 17 | 12 | **11** |
| **OMEGA** | 3,29% | 5,05% | -14,77% | -3,13% | 5 | 6 | 5 | 9 | **1** |
| **FT DEF.** | 2,68% | 4,81% | -13,77% | -3,06% | 20 | 4 | 4 | 3 | **3** |
| **FT CONS.** | 3,41% | 5,23% | -15,54% | -3,24% | 2 | 14 | 10 | 13 | **7** |
| **FT GROWTH** | 3,50% | 5,33% | -15,88% | -3,28% | 1 | 15 | 12 | 15 | **9** |
| **FT AGGR.** | 3,17% | 5,21% | -14,77% | -3,30% | 9 | 12 | 5 | 16 | **8** |
| **VR** | 3,26% | 5,12% | -15,88% | -3,13% | 7 | 11 | 12 | 8 | **6** |
| **VARR** | 3,38% | 6,01% | -15,97% | -3,91% | 3 | 22 | 14 | 21 | **20** |
| **STARR** | 3,20% | 5,36% | -16,19% | -3,24% | 8 | 16 | 18 | 14 | **19** |
| **GRR RISK NEUT.** | 2,96% | 5,73% | -13,29% | -3,79% | 12 | 18 | 1 | 18 | **14** |
| **GRR RISK AV.** | 2,96% | 5,73% | -13,29% | -3,79% | 12 | 18 | 1 | 18 | **14** |
| **GRR RISK LOV.** | 2,96% | 5,73% | -13,29% | -3,79% | 12 | 18 | 1 | 18 | **14** |
| **MRAR 2** | 2,18% | 8,08% | -31,69% | -5,87% | 27 | 27 | 27 | 28 | **29** |
| **MRAR 20** | 2,86% | 6,16% | -21,51% | -3,93% | 16 | 23 | 22 | 22 | **23** |
| **MRAR 50** | 2,64% | 4,80% | -16,52% | -3,12% | 21 | 3 | 20 | 7 | **17** |
| **EW** | 3,29% | 6,63% | -26,20% | -4,65% | 6 | 25 | 26 | 24 | **22** |
| **COMPOSITE** | 2,87% | 5,06% | -16,48% | -3,20% | 15 | 7 | 19 | 11 | **18** |

Source: own elaboration

***Table 36. Rankings of portfolio series - Summary***

| | P=2 | P=4 | P=8 | P=2 | P=4 | P=8 | P=2 | P=4 | P=8 | P=2 | P=4 | P=8 | P=2 | P=4 | P=8 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **RANK RETURN** | **RANK RETURN** | **RANK RETURN** | **RANK VOLAT.** | **RANK VOLAT.** | **RANK VOLAT.** | **RANK DRAWD.** | **RANK DRAWD.** | **RANK DRAWD.** | **RANK CVAR 95%** | **RANK CVAR 95%** | **RANK CVAR 95%** | **FINAL RANK** | **FINAL RANK** | **FINAL RANK** |
| **SHARPE** | 13 | 11 | 24 | 6 | 7 | 1 | 6 | 6 | 5 | 6 | 7 | 1 | **3** | **6** | **3** |
| **TREYNOR** | 28 | 22 | 29 | 28 | 1 | 21 | 28 | 1 | 23 | 28 | 1 | 23 | **28** | **5** | **24** |
| **ALPHA J** | 4 | 16 | 28 | 19 | 18 | 24 | 21 | 20 | 25 | 21 | 20 | 25 | **21** | **20** | **25** |
| **INF. RATIO** | 27 | 10 | 23 | 29 | 25 | 29 | 29 | 26 | 29 | 29 | 26 | 27 | **29** | **23** | **28** |
| **M2** | 13 | 11 | 24 | 6 | 7 | 1 | 6 | 6 | 5 | 6 | 7 | 1 | **3** | **6** | **3** |
| **ADJ. SHARPE** | 1 | 26 | 26 | 23 | 23 | 26 | 23 | 23 | 24 | 26 | 23 | 26 | **23** | **26** | **25** |
| **DIAMAN** | 2 | 2 | 22 | 20 | 26 | 28 | 19 | 24 | 28 | 14 | 25 | 29 | **18** | **22** | **27** |
| **CALMAR** | 13 | 11 | 10 | 6 | 7 | 17 | 6 | 6 | 21 | 6 | 7 | 17 | **3** | **6** | **21** |
| **STERLING** | 6 | 8 | 17 | 14 | 3 | 8 | 14 | 2 | 9 | 16 | 3 | 10 | **13** | **2** | **10** |
| **BURKE** | 13 | 7 | 18 | 6 | 2 | 9 | 6 | 2 | 15 | 6 | 2 | 5 | **3** | **1** | **12** |
| **MARTIN** | 13 | 8 | 18 | 6 | 3 | 9 | 6 | 2 | 15 | 6 | 3 | 5 | **3** | **2** | **12** |
| **SORTINO** | 13 | 17 | 11 | 6 | 11 | 5 | 6 | 6 | 10 | 6 | 7 | 4 | **3** | **10** | **2** |
| **K3** | 6 | 20 | 4 | 14 | 15 | 13 | 14 | 17 | 17 | 16 | 16 | 12 | **13** | **19** | **11** |
| **OMEGA** | 11 | 17 | 5 | 2 | 11 | 6 | 1 | 6 | 5 | 1 | 7 | 9 | **1** | **10** | **1** |
| **FT DEF.** | 13 | 11 | 20 | 6 | 7 | 4 | 6 | 6 | 4 | 6 | 7 | 3 | **3** | **6** | **3** |
| **FT CONS.** | 6 | 6 | 2 | 14 | 19 | 14 | 14 | 17 | 10 | 16 | 19 | 13 | **13** | **15** | **7** |
| **FT GROWTH** | 2 | 3 | 1 | 20 | 17 | 15 | 19 | 6 | 12 | 14 | 15 | 15 | **18** | **10** | **9** |
| **FT AGGR.** | 25 | 23 | 9 | 5 | 6 | 12 | 5 | 6 | 5 | 5 | 6 | 16 | **12** | **10** | **8** |
| **VR** | 6 | 5 | 7 | 14 | 5 | 11 | 14 | 2 | 12 | 16 | 5 | 8 | **13** | **4** | **6** |
| **VARR** | 24 | 1 | 3 | 13 | 20 | 22 | 13 | 16 | 14 | 13 | 18 | 21 | **20** | **14** | **20** |
| **STARR** | 6 | 15 | 8 | 14 | 16 | 16 | 14 | 17 | 18 | 16 | 17 | 14 | **13** | **18** | **19** |
| **GRR RISK NEUT.** | 21 | 27 | 12 | 24 | 27 | 18 | 24 | 27 | 1 | 23 | 27 | 18 | **24** | **27** | **14** |
| **GRR RISK AV.** | 21 | 27 | 12 | 24 | 27 | 18 | 24 | 27 | 1 | 23 | 27 | 18 | **24** | **27** | **14** |
| **GRR RISK LOV.** | 21 | 27 | 12 | 24 | 27 | 18 | 24 | 27 | 1 | 23 | 27 | 18 | **24** | **27** | **14** |
| **MRAR 2** | 26 | 4 | 27 | 27 | 24 | 27 | 27 | 24 | 27 | 27 | 24 | 28 | **27** | **21** | **29** |
| **MRAR 20** | 11 | 24 | 16 | 2 | 22 | 23 | 1 | 22 | 22 | 1 | 22 | 22 | **1** | **25** | **23** |
| **MRAR 50** | 29 | 19 | 21 | 1 | 14 | 3 | 4 | 14 | 20 | 3 | 14 | 7 | **11** | **15** | **17** |
| **EW** | 5 | 25 | 6 | 22 | 21 | 25 | 22 | 21 | 26 | 22 | 21 | 24 | **22** | **24** | **22** |
| **COMPOSITO** | 20 | 21 | 15 | 4 | 13 | 7 | 3 | 14 | 19 | 4 | 13 | 11 | **3** | **15** | **18** |

Source: own elaboration

## 4.5 Conclusions

In this chapter, we developed a methodology for the construction of **funds of funds portfolios** based on rolling estimates of the performance measures. They were used either individually or by combining them together to form the composite indicators.

The aim was to analyse certain **features** of the resulting funds of funds portfolios.

In order to examine how much the composition of the different portfolios varied over time, we calculated a **turnover measure**. In this way, we tried to evaluate the stability of both the performance measures and the composite indicators.

Finally, we compared the performances in terms of **risk and return** through statistics referred to the funds of funds.

The **results** did not show the presence of an indicator able to create portfolios among the best in terms of both risk and return. In fact, they revealed that the performance measures capable of creating portfolios with high returns often lead to selecting worse assets in terms of risks (volatility, maximum drawdown and CVaR 95%), and vice versa.

# CHAPTER FIVE

## 5 Conclusions

This study had the **aim** of thoroughly examining certain features of the most common performance measures.

In order to perform this analysis effectively, we started studying the **financial instruments** used for the first part of the examination. After a theoretical overview of the world of mutual funds, we described the selecting process of the **funds.** They were picked out according to the following characteristics: pure balanced funds distributed in Italy; one for each investment company; euro denominated; with at least 10 years of history. Those with an extreme value of asset under management were not involved in the process.
The benchmark was 50% composed of Bloomberg Barclays Euro Aggregate Bond Index and 50% of the FTSE All World Index; the risk-free rate was set to zero.
We obtained for each fund 179 monthly prices, equivalent to 178 monthly returns (from May 2002 to February 2017).
After an analysis of the returns distributions, we concluded that, except for single cases (FUND 7 for monthly returns, FUND8 for 60 months rolling returns), they could not be approximated by a normal distribution.

In the second part, the **performance measures** were described and grouped into five families: general risk measures, measures based on drawdowns, measures based on partial moments, measures based on quantiles and measures derived from utility functions.

Subsequently, the **correlations** and the **stability indicators** of each measure were analysed through an empirical study on the selected funds.

The analysis of the **within group correlation** was presented because some measures could be considered equivalent if highly correlated. The study was carried out using the Spearman rank correlation; if its value was greater than 0,915, then the analysed measures were considered "highly correlated". We firstly performed a **static analysis** of the rank correlation, and then we used a **rolling approach** in order to test the correlation between rankings over time.

The **results** of the entire study about correlation are described as follows:

***- Traditional and other unclassified performance measures:*** the information ratio, the adjusted Sharpe ratio, the Diaman ratio and the Jensens's alpha did **not** result highly correlated to the other measures. The M2 measure was perfectly correlated to the Sharpe ratio. The Sharpe ratio turned out to be highly correlated with the Treynor ratio.

***- Measures based on drawdown:*** they resulted in having a high within group correlation; in fact, all the indicators were found highly correlated to each other.

- ***Measures based on partial moments:*** the Sortino ratio, the K3 ratio, the Omega ratio and the Farinelli-Tibiletti for defensive investors proved to be concordant, thus highly correlated to each other. The FT for a conservative investor was highly correlated with the FT for a "growth" investor, whereas the FT for an aggressive investor was slightly correlated with all the other indices of the group.

- ***Measures based on quantiles:*** we found a high correlation between the VR ratio and the STARR ratio. The three GRR ratios resulted highly correlated with each other, whereas the VARR ratio showed low rank correlations with the other measures of the group.

***- Measures derived from utility functions:*** the MRAR 20 and MRAR 50 were found concordant, whereas the MRAR 2 was slightly correlated to the other measures of the group.

The next step consisted in calculating the **stability** of all the performance measures. The stability index pointed out the degree of similarity over time among rankings of funds induced by the different performance measures.

Many **results** emerged from this analysis:

- The ***case*** in which the sample was divided into ***8 sub-periods*** was the one with the most stable indices.
- The ***category*** of the ***measures derived from utility functions*** (MRAR) appeared to be one of the most stable, confirming what was found in Menardi & Lisi (2012a,b).
- The ***MRAR 50*** resulted the most stable ***measure***, whereas the ***Sharpe ratio adjusted for skewness and kurtosis*** turned out to be the less stable.
- The ***measures built for risk averse investors*** were, on average, more stable than those built for risk lovers.

In order to create a **composite indicator** which maximizes the stability over time we selected a restricted group of performance measures. For starters, we chose the performance measures uncorrelated with other indicators. Then, we selected the most stable measures among those correlated. In particular, the final indices were: the Sharpe ratio, the Jensen's alpha, the information ratio, the Diaman ratio, the Martin ratio, the K3 ratio, the FT ratio for conservative investors, the FT ratio for aggressive investors, the VR ratio, the VARR ratio, the GRR for risk averse investors, the MRAR 2 and the MRAR 50. The composite indicator was built as a linear combination of the abovementioned single performance measures, selecting the composition which **maximizes the stability** over time.

- When ***P equalled to two***, the composite indicator was not able to improve the stability of the most stable performance measure. However, many different compositions reached the same stability value of the most stable indicator, the information ratio.
- Also when ***P equalled to four*** the composite indicator was not able to improve the stability of the most stable performance measure. However, in 4 occasions the composite indicators reached the same stability values of the most stable indicator, the MRAR 50.
- It is only when ***P equalled to eight*** that the composite indicator had a gain in term of stability in comparison to the single measures.

The composite indicators, the single performance measures and an equally weighted portfolio were finally analysed through a **fund selection process**. We focused our study on 13 new funds, different from the ones previously used.

- From the analysis of the ***turnover rate***, the MRAR 50 was confirmed to be one of the most stable measures; moreover, the indicators built for risk averse investors were found, on average, more stable in comparison to those built for risk lover investors. With P=4 the composite indicator previously built proved to be the indicator with the lowest turnover rate. However, the results showed that a low turnover rate (high stability) did not always imply high returns.
- Analysing the ***return and risk aspects***, the results did not show the presence of an indicator able to create portfolios among the best in terms of both risk and returns; thus the best performance measures in terms of returns were often among the worst in terms of risk, and vice versa.

The answers to the **original questions** contained in the **Introduction** were all answered.

However, **further implementations** may be possible.

Future research could focus on applying the same methodology used in this study on a larger set of performance indicators. For example, Caporin, Jannin, Lisi & Maillet (2014) presented a description of "several dozen PMs collected in the academic literature over the last 50 years" which may be used for this additional study.

Furthermore, in order to make the results of this research more significant, the set of analysed mutual funds needs to be increased.

Lastly, the correlation and the stability may be studied isolating different contexts, for example, specific market conditions.

**Summarizing**, the study should help the investors in selecting the appropriate performance measure to evaluate an investment. Moreover, this research reinforces the idea that the choice of an indicator rather than another has some consequences because of the intrinsic characteristics of every single measure. The complete knowledge of these features is then crucial for properly selecting the financial instruments to invest in.

## References

Assogestioni. (2003). Guida alla classificazione.

Bachelier, L. (1900). *The theory of speculation* Gauthier-Villars.

Bacon, C. R. (2011). *Practical portfolio performance measurement and attribution* John Wiley & Sons.

Bernardi, D., & Bertelli, R. (2014). The Diaman ratio. *Wilmott, 2014*(71), 70-77.

Biglova, A., Ortobelli, S., Rachev, S. T., & Stoyanov, S. (2004). Different approaches to risk estimation in portfolio theory. *The Journal of Portfolio Management, 31*(1), 103-112.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637-654.

Bodson, L., Coën, A., & Hübner, G. (2008). How stable are the major performance measures? *Journal of Performance Measurement, 13*(1), 21-30.

Borsa Italiana. (2007). *Cosa sono i fondi comuni di investimento?* Retrieved May 15, 2017, from http://www.borsaitaliana.it/notizie/sotto-la-lente/fondiinvestimento.htm

Broner, F. A., Gelos, R. G., & Reinhart, C. M. (2006). When in peril, retrench: Testing the portfolio channel of contagion. *Journal of International Economics, 69*(1), 203-230.

Burghardt, G., & Walls, B. (2003). Understanding drawdowns. *Managed Futures for Institutional Investors: Analysis and Portfolio Construction*, 181-199.

Burke, G. (1994). A sharper Sharpe ratio. *Futures, 23*(3), 56.

Caporin, M., Jannin, G. M., Lisi, F., & Maillet, B. B. (2014). A survey on the four families of performance measures. *Journal of Economic Surveys, 28*(5), 917-942.

Caporin, M., & Lisi, F. (2011). Comparing and selecting performance measures using rank correlations.

Chakravorti, S., & Lall, M. S. (2004). *Managerial incentives and financial contagion (EPub)* International Monetary Fund.

Cogneau, P., & Hubner, G. (2009). The 101 ways to measure portfolio performance.

Collins, S., & Mack, P. (1997). The optimal amount of assets under management in the mutual fund industry. *Financial Analysts Journal, 53*(5), 67-73.

Damodaran, A. (2006). Investment valuation.

Dean, S., & Illowsky, B. (2017). Descriptive statistics: Skewness and the mean, median, and mode.

Di Bartolomeo, D. (2010). Strategy risk and the central paradox of active management.

Di Bartolomeo, D. (2014). Understanding the distribution of hedge fund returns.

Domian, D., Gibson, P., & Nanigian, D. (2015). Is your tax-managed fund manager hiding in the closet? *The Journal of Wealth Management, 18*(2), 67.

Eling, M., Farinelli, S., Rossello, D., & Tibiletti, L. (2011). One-size or tailor-made performance ratios for ranking hedge funds? *Journal of Derivatives & Hedge Funds, 16*(4), 267-277.

Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking & Finance, 31*(9), 2632-2647.

European Parliament and Council of the European Union. (2009). Directive 2009/65/EC of the european parliament and of the council. *Official Journal of the European Union.*

Farinelli, S., Ferreira, M., Rossello, D., Thoeny, M., & Tibiletti, L. (2009). Optimal asset allocation aid system: From "one-size" vs "tailor-made" performance ratio. *European Journal of Operational Research, 192*(1), 209-215.

Feibel, B. J. (2003). *Investment performance measurement* John Wiley & Sons.

Gemmill, G., Hwang, S., & Salmon, M. (2006). Performance measurement with loss aversion. *Journal of Asset Management, 7*(3-4), 190-207.

Gottardo, P., & Murgia, M. (1996). Variabili fondamentali e performance dei fondi comuni di investimento. *Rivista Internazionale Di Scienze Economiche e Commerciali, 43*(4), 927946.

Gregoriou, G. N., & Gueyie, J. (2003). Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio. *The Journal of Wealth Management, 6*(3), 77-83.

Grinold, R. C., & Kahn, R. N. (2000). Active portfolio management.

Haslem, J. A. (2010). *Mutual funds: Portfolio structures, analysis, management, and stewardship* John Wiley & Sons.

Hull, J. (2012). *Risk management and financial institutions, web site* John Wiley & Sons.

Ibert, M., Kaniel, R., Van Nieuwerburgh, S., & Vestman, R. (2017). Are mutual fund managers paid for investment skill? *The Review of Financial Studies, 31*(2), 715-772.

ICI. (2016). Characteristics of mutual fund investors. *ICI Research Perspective, 22*(7).

Indro, D. C., Jiang, C. X., Hu, M. Y., & Lee, W. Y. (1999). Mutual fund performance: Does fund size matter? *Financial Analysts Journal, 55*(3), 74-87.

International Monetary Fund. (2015). Chapter 3 - the asset management industry and financial stability. *Global Financial Stability Report - April 2015.*

Ivanovski, Z., Stojanovski, T., & Narasanov, Z. (2015). Volatility and kurtosis of daily stock returns at MSE. *UTMS Journal of Economics, 6*(2), 209-221.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance, 23*(2), 389-416.

Kaplan, P. D., & Knowles, J. A. (2004). Kappa: A generalized downside risk-adjusted performance measure. *Journal of Performance Measurement, 8*, 42-54.

Kenney, F., & Keeping, E. (1951). Mathematics of statistics-part two.

Kestner, L. N. (1996). Getting a handle on true performance. *Futures-Cedar Falls Iowa then Chicago-, 25,* 44-47.

Le Sourd, V. (2007). Performance measurement for traditional investment. *Financial Analysts Journal, 58*(4), 36-52.

Lehman, J., & Phelps, S. (2008). *West's encyclopedia of american law* (2nd ed.) Gale.

Lisi, F., & Caporin, M. (2012). On the role of risk in the Morningstar rating for mutual funds. *Quantitative Finance, 12*(10), 1477-1486.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, *36*, 394-419.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77-91.

Martin, P., & McCann, B. (1989). The investor's guide to fidelity funds.

Martin, R. D., Rachev, S. Z., & Siboulet, F. (2003). Phi-alpha optimal portfolios and extreme risk management. *The Best of Wilmott 1: Incorporating the Quantitative Finance Review, 1*, 223.

Menardi, G., & Lisi, F. (2012a). On the stability of performance measures over time: An empirical study. *Journal of Performance Measurement.*

Menardi, G., & Lisi, F. (2012b). Are performance measures equally stable? *Annals of Finance, 8*(4), 553-570.

Modigliani, F., & Modigliani, L. (1997). Risk-adjusted performance. *The Journal of Portfolio Management*, *23*(2), 45-54.

Morningstar. (2007). The Morningstar rating methodology. *Morningstar Methodology Paper.*

Pianca, P. (2003). Metodi quantitativi per la misurazione della performance dei fondi comuni di investimento.

Plastira, S. (2014). Performance evaluation of size, book-to-market and momentum portfolios. *Procedia Economics and Finance, 14*, 481-490.

Pozen, R., Hamacher, T., & Phillips, D. T. (2015). *The fund industry*, *website: How your money is managed* John Wiley & Sons.

Qian, E., & Hua, R. (2005). Active risk and information ratio. *The world of risk management* (pp. 151-167) World Scientific.

Rajan, R. G. (2006). Has finance made the world riskier? *European Financial Management, 12*(4), 499-533.

Rogo, B. (2013). La nuova normativa sui fondi d'investimento, la misurazione del rischio per la trasparenza d'informazione.

Sekhar, G. S. (2017). *The management of mutual funds* Springer.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425-442.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of Business, 39*(1), 119-138.

Sortino, F. A., & Satchell, S. (2001). *Managing downside risk in financial markets* Elsevier.

Sortino, F., & Van Der Meer, R. (1991). Downside risk. *The Journal of Portfolio Management, 17*(4), 27-31.

Tibiletti, L., & Farinelli, S. (2003). Upside and downside risk with a benchmark. *Atlantic Economic Journal, 31*(4), 387-387.

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*(1), 63-75.

Young, T. W. (1991). Calmar ratio: A smoother tool. *Futures, 20*(1), 40.

Zakamouline, V., & Koekebakker, S. (2009). Portfolio performance evaluation with generalized Sharpe ratios: Beyond the mean and variance. *Journal of Banking & Finance, 33*(7), 1242-1254.