La collana della Scuola di Alta Formazione Dottorale accoglie le migliori tesi di dottorato dell'Università degli Studi di Bergamo, insignite della dignità di stampa e sottoposte a procedura di *blind peer review*.

After a brief overview of the literature's framework and a methodological introduction, this thesis presents two different portfolio selection models. The first adopts a Markov bi-variate process to characterize the evolution of wealth and volatility generated by the portfolios. The second presents a new deviation measure based on quantile regression and it is used to develop a tracking error portfolio subjected to enhancement with second order stochastic constraints. After a theoretical overview, the problems are addressed formally and with empirical applications. Both the models out-perform the benchmark in empirical analysis making this work a promising starting point for new research development.

Marco Bonomelli obtained his Master's Degree in Management, Finance and International Business at the University of Bergamo in 2016 and his PhD in Analytics for Economics and Business at the University of Bergamo in 2020 (32nd cycle). Currently, he works for a consulting company in Milano and he keeps on some research collaborations. His research focuses on portfolio theory with a specific focus on portfolio optimization.

ISBN: 978-88-97413-50-9
DOI: 10.13122/978-88-97413-50-9

Marco Bonomelli MODELS AND METHODS FOR PORTFOLIO SELECTION

31

Collana della Scuola di Alta Formazione Dottorale
- 31 -

Marco Bonomelli

MODELS AND METHODS FOR PORTFOLIO SELECTION

UNIVERSITÀ DEGLI STUDI DI BERGAMO

2021

Collana della Scuola di Alta Formazione Dottorale

- 31 -

Collana della Scuola di Alta Formazione Dottorale

Diretta da Paolo Cesaretti

Ogni volume è sottoposto a *blind peer review*.

ISSN: 2611-9927

Sito web: https://aisberg.unibg.it/handle/10446/130100

Marco Bonomelli

MODELS AND METHODS FOR PORTFOLIO SELECTION

Università degli Studi di Bergamo

2021

Models and methods for portfolio selection /
Marco Bonomelli. – Bergamo :
Università degli Studi di Bergamo, 2018.
(Collana della Scuola di Alta Formazione Dottorale; 31)

ISBN: 978-88-97413-50-9
DOI: 10.13122/978-88-97413-50-9

Progetto grafico: Servizi Editoriali – Università degli Studi di Bergamo
© 2018 Università degli Studi di Bergamo
via Salvecchio, 19
24129 Bergamo
Cod. Fiscale 80004350163
P. IVA 01612800167

https://aisberg.unibg.it/handle/10446/200301

Acknowledgements

I would firstly like to thank my supervisor Prof. Rosella Giacometti for the great support, kindness and attention she has always demonstrated to me. Then, I would like to thank all the Professors of the department of management, economics and quantitative methods of the University of Bergamo. They have always been a source of inspiration and advice for my research. In particular, I want to express my gratitude to the coordinators of the PhD program, Prof. Marida Bertocchi and Prof. Adriana Gnudi.

I would like to thank my colleagues of the office 206, sharing with them this adventure has made it lighter. Here I have found people who have become more than friends to me. Thanks to Gilda for the kindness.

In the end, I have to be grateful to my Family and my Friends, their support has been crucial for these intense years. My parents and Marta have been by my side at every moment of my life, their closeness has been reassuring in the most important choices, even in the hardest ones.

To my Family

Table of contents

Introduction **1**

Chapter 1. Joint tails impact in stochastic volatility portfolio selection models **13**

1.1 Introduction 13
1.2 Approximating a stochastic volatility model with a bivariate Markov chain 15
1.3 Joint tails in portfolio selection stochastic volatility models 19
1.3.1 Tail portfolio performance measures 21
1.4 An empirical analysis of the stochastic volatility portfolio model . . 23
1.4.1 Portfolio selection comparison 25
1.4.2 On the impact of the portfolios optimized on the tails 28
1.5 Conclusions 31

Chapter 2. Enhanced tracking error quantile regression **33**

2.1 Introduction 33
2.2 Index tracking problem: a quantile regression approach 34
2.3 Stochastic dominance constraints 37
2.4 Enhanced Indexing Problem with Stochastic Dominance Constraints 39
2.5 Empirical Application 41
2.5.1 Portfolio wealth 42
2.5.2 Accumulated cumulative empirical functions 43
2.5.3 Portfolio concentration 43
2.5.4 Statistical analysis 45
2.6 Conclusions 45
2.7 Annex: Choosing the calibrating window 48

List of figures **51**

List of tables **53**

References **55**

Introduction

It has been almost seventy years since Harry Markowitz postulated the fundamentals of portfolio theory and mean-variance analysis (see Markowitz (1952)). He had an intuition: each security must be seen with respect to its contribution at the entire portfolio in terms of wealth and risk. According to Markowitz's perspective, a portfolio is said to be efficient if, given a desired return level, it has the lowest variance. The idea is based on the evidence that efficient portfolios lie on the frontier of the set of every feasible portfolios. This frontier represents the trade off between risk and reward. Investors choose portfolios at the tangency point between their iso-utility curves and the frontier. Since then, several authors have tried to develop new theories and improve Markowitz's ideas and portfolio theory. Various mathematical methods are used to develop optimal portfolio allocation models. Concepts like Markov process and Stochastic dominance are here pivotal and they will be described in the following paragraphs.

Markov Process

In 1906 Andrei Andreevich Markov identified a particular sequence (later Markovian chain) of quantities in which future depends on the sequence of past events but it is independent of their order. Therefore the subsequent evolution of a random variable depends only on the present state. He stated this in opposition to the idea of Nekrasow according to which the random addends' independence is a necessary condition for the weak law of large numbers (see Seneta, 1996). In economics and finance, several authors have used Markov Chains to approximate processes of random variables. For instance, Tauchen (1986) has applied this stochastic model to approximate a continuous-valued auto-regression. The aim of the author was to provide a methodology able to find numerical solution to integral equations for those problems with discrete state spaces. D'Amico and Di Biase (2009) used a Markovian approach in order to identify changes in the composition of population without the presence of economic shocks. Their model was built as a semi-Markov chain.

In discrete time and continuous space (see Ibe, 2014), a Markov Process is defined as a stochastic process $\{X(t),\ t \in T\}$ where, the conditional CDF of $X(t_n)$ for

$$X(t_0), X(t_1), \ldots, X(t_{n-1})$$

with $t_0 < t_1 < \cdots < t_n$ depends only on $X(t_{n-1})$. Formally:

$$P[X(t_n) \leq x_n | X(t_{n-1}) \leq x_{n-1}, X(t_{n-2}) \leq x_{n-2}, \ldots, X(t_0) \leq x_0] =$$

$$= P[X(t_n) \leq x_n | X(t_{n-1}) \leq x_{n-1}]$$

Markov processes can be easily represented as trees. Bean, Kontoleon, and Taylor (2008) define such a structure "Markov Tree" and they represent it as a continuous-time Markovian Multitype Branching Process. Such kind of process admits the possibility that the states of the Markov chain may evolve and even extinguish. In their work two algorithms are presented to compute the extinction probability of the tree.

In this work Markov processes are used in the bivariate framework for jointly modeling the process of portfolio's wealth and variance. Such kind of processes are indicated with $W_t = (W_{x,t}, W_{\sigma,t})$, and they can be seen as follows: $W_{x,t}$ is the process generated by the portfolio return β and $W_{\sigma,t} = \exp \sigma_{X_{\beta,t}}$ is the one generated by the volatility. N states for the returns of portfolio and M states for its volatility are considered. Then, the support of Markov process, based on a set past values, $L \in \mathbb{N}^+$, is discretized. The range of the portfolio process is

$$\left(\min_{s=-L,\dots,0} W_{x,s}, \max_{s=-L,\dots,0} W_{x,s} \right) \times \left(\min_{s=-L,\dots,0} W_{\sigma,s}, \max_{s=-L,\dots,0} W_{\sigma,s} \right)$$

and it is divided into $N \cdot M$ bi-dimensional intervals $(a_i, a_{i-1}) \times (b_j, b_{j-1})$, where $\{a_i\}$ and $\{b_j\}$ are given by the two following decreasing sequences

$$\begin{aligned} a_i &:= u_x^i \max_s W_{x,s}, \quad i = 0, \dots, N \\ b_j &:= u_\sigma^j \max_s W_{\sigma,s}, \quad j = 0, \dots, M \end{aligned} \tag{1}$$

and $u_x := \left(\frac{\min_s W_{x,s}}{\max_s W_{x,s}} \right)^{1/N}$, $u_\sigma := \left(\frac{\min_s W_{\sigma,s}}{\max_s W_{\sigma,s}} \right)^{1/M}$ are two step factors, each representing the common ratio of a geometric progression.

The values of Markov process belonging to the bi-dimensional interval $(a_{i_x}, a_{i_x-1}) \times (b_{i_\sigma}, b_{i_\sigma-1})$ are approximated by the geometric mean indicated as the state $z^{(i)} = (z_x^{(i_x)}, z_\sigma^{(i_\sigma)})$ of the Markov chain

$$\begin{aligned} w_x^{(i_x)} &= \sqrt{a_{i_x} a_{i_x-1}} = u_x^{\frac{1-2i_x}{2}} \max_s W_{x,s}, \quad i_x = 1, \dots, N \\ w_\sigma^{(i_\sigma)} &= \sqrt{b_{i_\sigma} b_{i_\sigma-1}} = u_\sigma^{\frac{1-2i_\sigma}{2}} \max_s W_{\sigma,s}, \quad i_\sigma = 1, \dots, M. \end{aligned} \tag{2}$$

Thus, according to (1) and (2), $w_x^{(i_x)} = w_x^{(1)} u_x^{1-i_x}$ and $w_\sigma^{(i_\sigma)} = w_\sigma^{(1)} u_\sigma^{1-i_\sigma}$. Clearly, the M states for the volatility implicitly determine the relative M states of the portfolio volatility given by $\sigma_{(i_\sigma)} = \ln(w_\sigma^{(i_\sigma)})$, $i_\sigma = 1, \dots, M$.
The joint probability

$$\pi_i = \Pr\left(W_{x,t} \in (a_{i_x}, a_{i_x-1}), W_{\sigma,t} \in (b_{i_\sigma}, b_{i_\sigma-1}) \right)$$

of the portfolio return and its volatility to be respectively in states $w_x^{(i_x)}$ and $w_\sigma^{(i_\sigma)}$ is approximated by the number of times the process W_t is in the state $z^{(i)} = (w_x^{(i_x)}, w_\sigma^{(i_\sigma)})$ divided by the total number of joint observations. As consequence of the homogeneous property of Markov chain, the transition matrix is constant over the time and is given by $\Pi = [\pi_{ij}]_{i,j\in I}$, where:

$$\pi_{ij} = \Pr\left(W_{\tau+1} = w^{(j)} | W_\tau = w^{(i)}\right),\ i,j \in I$$

is the probability to move from the generic state $w^{(i)}$ to $w^{(j)}$ in one period of time. These transition probabilities can be seen in a transition matrix form as follows:

$$\Pi_{W_x,W_\sigma} = \begin{bmatrix} \pi_{11,11} & \pi_{11,21} & \cdots & \pi_{11,m1} & \pi_{11,12} & \cdots & \pi_{11,mn} \\ \pi_{21,11} & \pi_{21,21} & \cdots & \pi_{21,m1} & \pi_{21,12} & \cdots & \pi_{21,mn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \pi_{m1,11} & \pi_{m1,21} & \cdots & \pi_{m1,m1} & \pi_{m1,12} & \cdots & \pi_{m1,mn} \\ \pi_{12,11} & \pi_{12,21} & \cdots & \pi_{12,m1} & \pi_{12,12} & \cdots & \pi_{11,mn} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \pi_{mn,11} & \pi_{mn,21} & \cdots & \pi_{mn,m1} & \pi_{mn,12} & \cdots & \pi_{mn,mn} \end{bmatrix}$$

In particular, the matrix Π_{W_x,W_σ} is a stochastic matrix where the sum of every row is equal to one. The entry $\pi_{2i,j1}$ represents the probability for the portfolio's return process W_x to reach the state j from the actual state 2 and for the volatility process W_σ to reach the state 1 given it is now in i.

Duan and Simonato (2001) proposed a methodology for evaluating American options in a GARCH framework. They presented a bivariate Markov process for the evolution of volatility and price. In their work the authors show that the price computed via the approximation of the Markov chain converges to the theoretical price and to the target GARCH process. Using such method it is possible to discretize the evolution and therefore to simplify the pricing process since the expected value is computed as a simple product. Duan and Simonato (2001) proved that assuming an infinite number of states, a Markov chain reproduces exactly the probabilistic behavior of the target GARCH. The transition matrix in this model appears to be highly sparse thanks to the feature of the GARCH process. Sparse matrices are extremely useful for optimization, because they allow to avoid several operations thanks to the possibility to represent them through the indices of the non-zero elements only. This feature allows to treat also large problems in a quite easy way. The computational complexity is one of the main problems encountered in evaluating American options due to the possibility of exercise before the maturity.

Ortobelli Lozza and Iaquinta (2008) propose a markovian model for pricing European and

American options. In particular they propose a non-parametric approach for pricing contingent claims. They compare their results with the traditional Black and Scholes model.

Stochastic Dominance

Despite its elegance the Mean-Variance Efficient Frontier has some limitations. Firstly, the model has a limited validity and cannot be generalized; it is proved that assumptions are correct under an elliptical probability distribution or a quadratic utility function (see Bawa, 1975). Secondly, many authors have pointed out that the risk-averse assumption is not sufficient for describing all the investors' preferences (see for example Stiglitz (1970), Pratt (1978)). In order to overcome such limitations, new models based on Von Neumann, Morgenstern, and Kuhn (2007) have been developed. According to this theory, investors take decisions with respect to an utility comparison. Given a random vector of returns $R = [r_1, r_2, \ldots, r_N]$, $\alpha \in R^N$ and $\beta \in R^N$. A portfolio αR is preferred over a βR if the expected utility is higher: $\mathbb{E}[U(\alpha R)] \geq \mathbb{E}[U(\beta R)]$. Theories in this field admit a generalization, are more flexible but they collide with the lack of knowledge of the true investors' utility function. To overcome this point, researchers look for methods to order investors' choice with respect to homogeneous class of risk attitude; stochastic dominance based models are part of this family. Stochastic dominance is a partial order among random variables. It is said to be partial because there could exist some elements of the set for which it is non possible to formulate an ordering. There are different orderings based on stochastic dominance. In portfolio theory, stochastic dominance rules have been used to justify the reward-risk approaches. Despite stochastic dominance have been introduced by Karamata (1932), it made its appearance in financial research between 1969 and 1970 by Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970) and Whitmore (1970).
A complete characterization of stochastic dominance orders can be found in Whang (2019). Let X and Y be two random variable with $\mathcal{F}_X$ and $\mathcal{F}_Y$ the respective Cumulative Density Function (CDF).
$Q_k(\tau) = \inf\{x : \mathcal{F}_k(x) \geq \tau\}$ denote the quantile function of the distribution. The stochastic dominance for different orders can be defined as follows:

1. First order (FSD): Let $\mathcal{U}_1$ be the set of monotone non-decreasing utility functions. The random variable X is said to stochastically dominate the random variable Y at the first order, $X \overset{FSD}{\succeq} Y$, if one of the following equivalent conditions holds:

   (a) $\mathbb{E}[u(X)] - \mathbb{E}[u(Y)] \geq 0 \;\; \forall \;\; u \in \mathcal{U}_1$

   (b) $\mathcal{F}_Y(z) - \mathcal{F}_X(z) \geq 0 \;\; \forall \;\; z \in \mathbb{R}$

   (c) $Q_X(\tau) - Q_Y(\tau) \geq 0 \;\; \forall \;\; \tau \in [0, 1]$

A choice made according to FSD is accepted by insatiable investors. Jarrow (1986) proves that, under some conditions, FSD is a sufficient condition for the presence of arbitrage opportunities.

2. Second order (SSD): Let define $\mathcal{U}_2$ the set of monotone non decreasing and concave utility functions. The random variable X is said to stochastically dominate the random variable Y at the second order, $X \overset{SSD}{\succeq} Y$, if one of the following equivalent conditions holds:

   (a) $\mathbb{E}[u(X)] - \mathbb{E}[u(Y)] \geq 0 \;\; \forall \;\; u \in \mathcal{U}_2$

   (b) $\int_{-\infty}^{z} [\mathcal{F}_Y(t) - \mathcal{F}_X(t)] \, dt \geq 0 \;\; \forall \;\; z \in \mathbb{R}$

   (c) $\int_0^{\tau} [Q_X(p) - Q_Y(p)] \, dp \geq 0 \;\; \forall \;\; \tau \in [0,1]$

   A choice made according to SSD is accepted by insatiable and risk averse investors.

3. nth order: Let define $\mathcal{U}_n = \left\{u(\cdot) : u' \geq 0, u'' \leq 0, \ldots, (-1)^{s+1} u^{(s)} \geq 0\right\}$. The random variable X is said to stochastically dominate the random variable Y at the n-th order, $X \overset{NSD}{\succeq} Y$, if one of the following equivalent conditions holds:

   (a) $\mathbb{E}[u(X)] - \mathbb{E}[u(Y)] \geq 0 \;\; \forall \;\; u \in \mathcal{U}_n$

   (b) $\left[\mathcal{F}_Y^{(s)}(z) - \mathcal{F}_X^{(s)}(z)\right] \geq 0 \;\; \forall z \in \mathbb{R}$ and
   $\left[\mathcal{F}_Y^{(r)}(\infty) - \mathcal{F}_X^{(r)}(\infty)\right] \geq 0 \;\; \forall \;\; r \in [1, \ldots, s-1]$

   (c) $[Q_X^s(p) - Q_Y^s(p)] \, dp \geq 0 \;\; \forall \;\; \tau \in [0,1]$ and
   $[Q_X^r(1) - Q_Y^r(1)] \, dp \geq 0 \;\; \forall \;\; r \in [1, \ldots, s-1]$

The definitions above are related to the concept of weak stochastic dominance. If the inequalities become strict in some case, the stochastic dominance is said strong.

Stochastic dominance constraints' formulation

The introduction of Stochastic Dominance constraints in finance has been widely analyzed in literature. The most studied Stochastic Dominance order is the second (SSD) for its relevance in investors' preferences. Several authors have proposed different methods to include dominance constraints in optimization problems. In the following the most widespread models are presented.

Roman, Mitra, and Zverovich (2013) present a model for an index tracking error with enhancement based on SSD, their work is based on the model presented in Fábián, Mitra, and

Roman (2011), where the authors developed a cutting-plane method in order to make the problem tractable from a computational point of view. In their results, the proposed model is able to return outperforming portfolios with respect to the benchmark.
Recently, authors have presented models with the aim to linearize the stochastic dominance constraints. Here two different approaches for the SSD models are briefly presented. The stochastic dominance relation and relative constraints could be expressed in linear representation, as in Dentcheva and Ruszczyński (2006), introducing slack variables $s_{i,t}$ representing shortfall of $R^t\beta$ below y_i in realization t for $t = 1, \dots, T$:

$$\begin{aligned} &\sum_{n=1}^{N} R_{t,n}\beta_n + s_{i,t} \geq y_i \quad && \forall i = 1, \dots, T; \;\; \forall t = 1, \dots, T \\ &\sum_{t=1}^{T} s_{i,t} \leq \mathbb{E}[(y_i - Y)_+] && \forall i = 1, \dots, T \\ &s_{i,t} \geq 0 && \forall i = 1, \dots, T; \;\; \forall t = 1, \dots, T \end{aligned} \tag{3}$$

Differently, Kuosmanen (2004) and Kopa (2010) propose another linear formulation of the second order stochastic dominance based on the presence of a double stochastic permutation matrix $Z_{T\times T} = \{z_{r,c}\}$ (each element is positive and the sums of each row and each column are equal to one). If any additional assumptions about entries are made, the model has SSD constraint and it is linear, but requiring $z_{r,c} \in \{0,1\}$, the constraint become FSD, and the resulting problem is Mixed Integer.
Let us assume that the returns have a discrete joint distribution, a second order stochastic dominance $X \overset{SSD}{\succeq} Y$ constraint is satisfied if and only if there exists a double stochastic matrix Z, with $z_{r,c} \in [0,1]$ where $X = R\beta$ is the vector of portfolio returns to be tested and Y is the benchmark portfolio, as such:

$$\begin{aligned} &X \geq ZY \\ &\sum_{r=1}^{T} z_{r,c} = 1 \quad \forall c = 1, \dots, T \\ &\sum_{c=1}^{T} z_{r,c} = 1 \quad \forall r = 1, \dots, T \\ &0 \leq z_{r,c} \leq 1 \quad \forall r = 1, \dots, T; \;\; \forall c = 1, \dots, T \end{aligned} \tag{4}$$

The stochastic dominance constraint can be very restrictive and it can reduce the feasible set to an empty one. For this reason several authors tested some relaxations of the constraint. Leshno and Levy (2002) introduced, for the first and second order, the concept of *Almost Stochastic Dominance* (ASD) with the following definition: Let X and Y be two random variables, and

F and G denote their CDFs. For $\mathbb{E}_F(X) \geq \mathbb{E}_G(X)$ and for $\epsilon < 0.5$ as suggested by the authors. The Almost First Stochastic Dominance Order (AFSD) is defined as follows:
F dominates G by $\epsilon-$Almost FSD $F \overset{Almost(\epsilon)}{\succeq_1} G$ if and only if, $\epsilon \int_{S_1}[F(t) - G(t)]dt \leq \|F - G\|$;
The Almost Second Stochastic Dominance Order (ASSD) is defined as follows:
F dominates G by $\epsilon-$Almost ASD $F \overset{Almost(\epsilon)}{\succeq_2} G$ if and only if, $\epsilon \int_{S_2}[F(t) - G(t)]dt \leq \|F - G\|$;
Where $\|F - G\| = \int_0^1 |F(t) - G(t)| \, dt$ and the left side of the inequalities represents the amount of "correction" for making the stochastic dominance assumptions valid.

Testing stochastic dominance

The literature with respect stochastic dominance's test is wide. It is generally accepted a first classification about the assumptions underlying the test as the one presented in Whang (2019):

(A.) $H_0 \quad F_1^{(S)}(x) \leq F_2(x)^{(S)} \; \forall \; x$ versus $H_1 \quad F_1(x)^{(S)} > F_2(x)^{(S)}$ for some x;

(B.) $H_0 \quad F_1(x)^{(S)} \geq F_2(x)^{(S)} \; \forall \; x$ versus $H_1 \quad F_1(x)^{(S)} < F_2(x)^{(S)}$ for some x;

(C.) $H_0 \quad F_1(x)^{(S)} = F_2(x)^{(S)} \; \forall \; x$ versus $H_1 \quad F_1(x)^{(S)} < F_2(x)^{(S)}$ for some x;

where S represents the stochastic dominance's order. The most common type of test is the first, where the null hypothesis consists in assuming the presence of stochastic dominance and the alternative hypothesis indicates the presence of some violations. Tests in this class are divided with respect to the methodology used for computing:

1. Comparing the CDFs at a finite number of grid points: these tests are based on easy computing method and they give adequate information about the point of violation. Despite these positive aspects, these tests ar not consistent with respect to the alternative hypothesis (see Rao et al. (1973)). Examples of such kind of tests are, among others, Anderson (1996) and Davidson and Duclos (2000).

2. Comparing the CDFs at all points inside an interval. Tests of this type are consistent with respect to all the hypothesis. Examples are the ones based on Kolmogorov-Smirnoff statistic (for example Barrett and Donald (2003) or the ones based on quantiles (see among the others Koenker and Xiao (2002))).

3. Tests non-classifiable with the definitions above. For example Robertson and Wright (1981) built a test based on the likelihood for discrete distributions and Schmid and Trede (1996) used a Mann?Whitney?Wilcoxon type test.

One of the most used test is the one in Barrett and Donald (2003) for its computational simplicity. This is constructed under the assumption of equal support for F_1 and F_2 where they are continuous functions. Let $X_{1,i}$ and $X_{2,i}$ be two independent random samples from F_1 and F_2. The test-statistic is computed as follows:

$$BD = \sqrt{\frac{NM}{N+M}} \sup_{x \in \mathbb{X}} \left(\overline{D}_{1,2}^{(S)}(x) \right)$$

where N and M are the two sample length and $\overline{D}_{1,2}^{(S)}(x)$ is the difference between the integrated empirical CDFs.

Deviation measures

A metric defines the distance between the elements of a given set, it must satisfy a set of axioms. As in Rachev (1991), a probability metric can be any functional suitable to measure the distance between two random quantities. Deviation measures are sub-class of the dispersion measures. In Stoyanov et al. (2008), the authors present the new class of relative deviation metrics able to satisfy the following properties:

P1. $\nu(X,Y) \geq 0$ and $\nu(X,X) = 0$;

P2. $\nu(X,Y) = \nu(Y,X) \ \ \forall X,Y$

P3. $\nu(X,Y) \leq \nu(X,Z) + \nu(Z,Y) \forall X,Y,Z$;

P4. $\nu(aX,aY) = a^s \nu(X,Y) \forall X,Y,a,s \geq 0$;

P5. $\nu(X+Z,Y+Z) = \nu(X,Y) \forall X,Y,Z$

P6. $\nu(X+c_1,Y+c_2) = \nu(X,Y) \forall X,Y$ and constants c_1,c_2;

If P1, P3, P4, P5 and P6 hold simultaneously, the metric is called of relative deviation. Stoyanov, Rachev, and Fabozzi (2008) proved that such kind of functional ν can lead to a deviation measure in the sense of Rockafellar, Uryasev, and Zabarankin (2002). The general deviation measures, whose properties are:

1. $D(Z+C) = D(Z)$ for all Z and constants C

2. $D(0) = 0$ and $D(\lambda Z) = \lambda D(Z)$ for all Z and all $\lambda > 0$

3. $D(Z,Z') \leq D(Z) + D(Z')$ for all Z and Z'

4. $D(Z) \geq 0$ for all Z, with $D(Z) > 0$ for non-constant Z

This kind of measures can be applied to tracking problems if $Z = X - Y$, then Y represents a benchmark and X in an investing portfolio. So, the minimization of this deviation measure could be seen as the solution of a tracking error problem. On the wave of this work, Rockafellar and Uryasev (2013) presented the relationship between deviation measures and risk measures. They proved the relation between a deviation measure constructed from the quantile regression and the Conditional Value at Risk (CVaR).

The search for a proper measure of risk is one of the main tasks in modern finance's literature. The cornerstones of this field are the Coherent Measures of Risk presented in Artzner et al. (1999). The authors state the properties a measure should satisfy to be coherent. Authors have shown how the use of a measure $\rho(\cdot)$ belonging to this family allows a correct risk evaluation. The properties proposed are:

1. Subadditivity: $\rho(X+Y) \leq \rho(X) + \rho(Y)$

2. Positive Homogeneity: $\rho(\lambda X) = \lambda\rho(X) \;\; for \;\; \lambda \geq 0 \; and \; \lambda \in \mathbb{R}$

3. Monotonicity: $\rho(X) \geq \rho(Y) \;\; if X \leq Y$

4. Translation Invariance: $\rho(X+m) = \rho(X) - m \;\; \forall \;\; m \in \mathbb{R}$

Tracking error

The tracking error (TE) is a measure of the distance between the returns of a portfolio and a benchmark (for example an index). The magnitude of these errors can be evaluated with several deviation measures $\sigma(\cdot)$. The lower is the measure, the better is the adherence of the investment portfolio to the trajectory of the replicating index.

Let be Y a random variable representing the benchmark and $R = [r_1, r_2, \ldots, r_N]$ a random vector representing the log returns of N assets. The investor choice is determined by the possibility to combine different assets in a portfolio. Consider $\beta \in \mathbb{R}^N$ be the vector of the weights and $X = R\beta$ is the relative portfolio. Now consider T equiprobable scenarios, where $R^t = [r_1^t, r_2^t, \ldots, r_N^t]$ with $t \in [1, \ldots, T]$ is a particular scenario of the random vector R and y^t with $t \in [1, \ldots, T]$ is a particular one of the random variable Y. So, the tracking error (TE) is the random vector $\varepsilon = R\beta - Y$, with $\varepsilon \in \mathbb{R}^T$.

There are several ways to build an index tracking portfolio since portfolio managers have different constraints and restrictions. A general benchmark tracking problem, without short-selling, can be defined as follows:

$$\begin{aligned}
\min_{\beta} \quad & \sigma\left(X-Y\right) \\
\text{s.t.} \quad & \sum_{n=1}^{N}\beta_n = 1 \\
& \mathbb{E}\left[X\right]-\mathbb{E}[Y] \geq K^* \\
& lb \leq \beta_n \leq ub \qquad \forall n=1,\ldots,N \\
& \beta_n \geq 0 \qquad \forall n=1,\ldots,N
\end{aligned} \tag{5}$$

where σ is the selected dispersion measure (see e. g. Stoyanov, Rachev, and Fabozzi, 2008). The first constraints represent the obligation to invest all the available wealth. The second constraint defines the minimum enhancement K^* of the portfolio's returns with respect to the benchmark's. The third constraint sets an upper and a lower bound of each weight. Then the fourth is the prohibition of short selling. In the financial literature three measures are broadly used in the this problem: the tracking error mean absolute deviation (TEMAD), the tracking error downside mean semi-deviation (TEDMS) and the tracking error volatility (TEV).

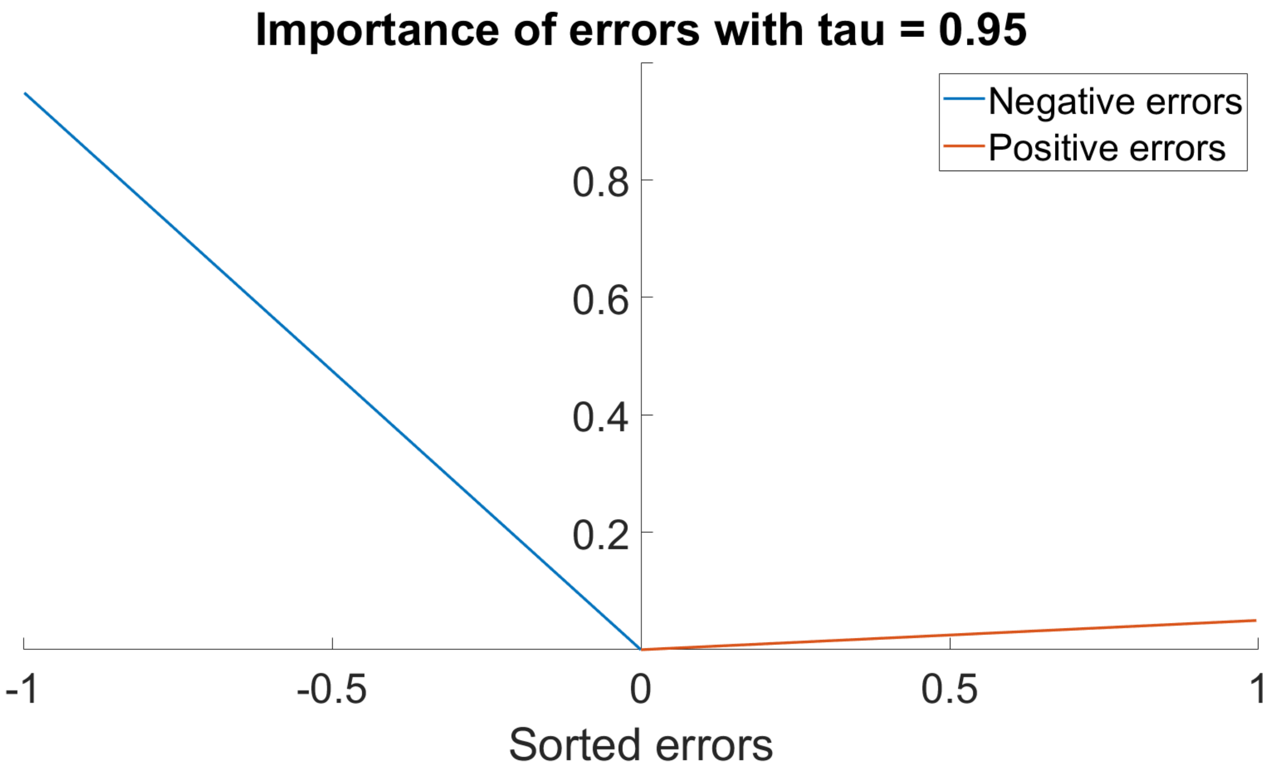

Figure 1: Interpretation of the relevance given to errors in quantile regression with $\tau = 0.95$.

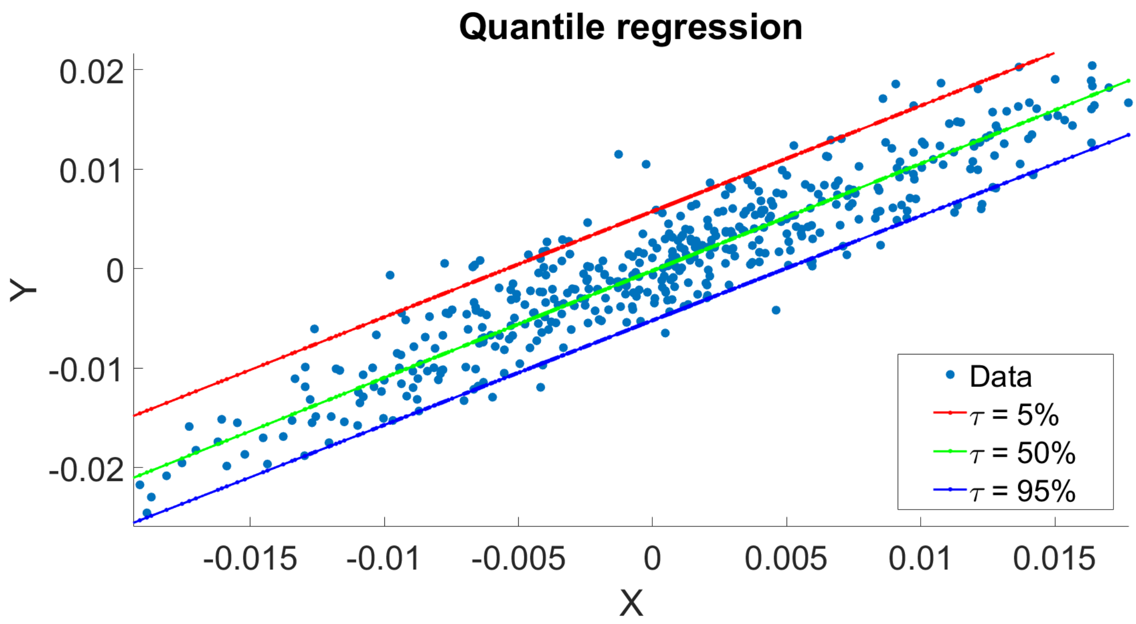

Figure 2: Interpretation of the quantile regression line with $\tau = 0.95, \tau = 0.5, \tau = 0.05$.

Quantile regression

Quantile regression (later called QR) has been introduced by Koenker and Bassett Jr (1978) with the aim to identify a new class of estimators. This methodology is more suitable to model phenomena characterized by heavy tails and allows a better adherence to the real world without any assumptions on distribution's parameters. Quantile regression is computationally simple, and provides a possible linking to a tracking error problem and it is easy to interpret with respect to risk management. As illustrated in Figure 1, QR allows to give different importance to errors depending on some features: for example, in order to contain risk of losses, it is necessary to set $\tau > 0.5$ to give heavier weights on negative errors. In figure 2 three regression line for different $\tau = [0.05, 0.5, 0.95]$ are shown. In the original formulation of 1978, quantile regression applied to index tracking takes the following formulation:

$$\min_{\beta \in \mathbb{R}^K} \left[\sum_{t \in \{t: y_t \geq x_t\beta\}} \tau \, |y_t - x_t\beta| + \sum_{t \in \{t: y_t < x_t\beta\}} (1-\tau) \, |y_t - x_t\beta| \right]$$

When the chosen quantile is the median, i.e. $\tau = 0.5$, the problem is reduced to least absolute error estimator.

In Rockafellar and Uryasev (2013) the authors present the risk quadrangle and, thanks to that, they state the connection between risk measures and deviation measures. In particular, they pay attention to quantile regression. They show how, in the optimization of the quantile regression, if the intercept ξ is considered in the model, it can be seen as the VaR at the τ quantile, with $\xi \; in \; \mathbb{R}$.

$$\min_{\beta,\xi\in\mathbb{R}^K}\left[\sum_{t\in\{t:\epsilon_t\geq\xi\}}\theta\left|\epsilon_t-\xi\right|+\sum_{t\in\{t:\epsilon_t<\xi\}}(1-\theta)\left|\epsilon_t-\xi\right|\right]$$

With respect to the previous formulation of the quantile regression, the error component $\epsilon_t = y_t - x_t\beta$ is introduced and the new value ξ cold provide useful information for about the risk level of portfolios.

The thesis is organized as follows. In the second chapter it is examined the impact of the joint tails of the portfolio return and its empirical volatility on the optimal portfolio choices. The portfolio return and its volatility dynamic are approximated by a bi-variate Markov chain constructed on its historical distribution. This allows the introduction of a non parametric stochastic volatility portfolio model without the explicit use of a GARCH type or other parametric stochastic volatility models. It is described the bi-dimensional tree structure of the Markov chain and it is discussed alternative portfolio strategies based on the maximization of the Sharpe ratio and of a modified Sharpe ratio that takes into account the behavior of a market benchmark. Then the impact of the portfolio and its stochastic volatility joint tails is empirically evaluated on optimal portfolio choices. In particular, it is presented the comparison of the out of sample wealth obtained optimizing the portfolio performances conditioned on the joint tails of the proposed stochastic volatility model. In the third chapter The construction of an enhanced index tracking portfolio with stochastic dominance constraint is investigated. It is discussed and compared to a general framework of the literature, then it is proposed an optimization model. The tracking error problem is dealt with by introducing a new deviation measure based on the use of quantile regression. The portfolio optimization problem is brought back to a linear formulation and it is improved with a linearization of the second order stochastic constraint present in literature. In order to define the optimal window for calibrating the model it has been conducted a sensitivity analysis. It is shown that in the out of sample framework, the built portfolios preserve a second order stochastic dominance with respect to the benchmark.

1. Joint tails impact in stochastic volatility portfolio selection models

1.1 Introduction

There is a general consensus (Engle (1982), Bollerslev (1986)) that the variance of the financial asset returns is time variant and a great amount of efforts are directing to realize mathematical models which, by choosing the variance dynamics as the model corner-stone, should be effectively able to model financial prices. Surely the GARCH model is a reference instrument to study the volatility dynamics, and among its advantages there is its high flexibility to be suitable to capture the most important features of the financial variables. As Glosten, Jagannathan, and Runkle (1993), and Nelson (1991) explain many GARCH type models and in particular the GARCH(1,1) model can be represented as a bivariate Markovian system (i.e., the state of the process is uniquely represented by price and variance states). This feature allows to approximate GARCH type models by a discrete Markov chain. The Markovian and semi-Markovian models have been used in different fields of the financial literature typically in option pricing and credit risk (see, among others, Duan and Simonato (2001) ; D'Amico and Di Biase (2009) , D'Amico et al. (2009) , D'Amico et al. (2010)), and in portfolio theory (see Angelelli and Ortobelli Lozza (2009), Ortobelli Lozza and Iaquinta (2008)). Elliott and Siu (2010) and Çanakoğlu and Özekici (2009) model the economic phases as a discrete Markov chain.

Duan and Simonato (2001) proposed a methodology based on a Markov chain process to approximate the asset price distribution and its conditional volatility under the risk-neutralized pricing measure when asset returns and its conditional volatility are modeled with a GARCH(1,1) model with Gaussian innovations.

In this line of research, this chapter investigate the construction of a non parametric Markov chain process that allows us to model the evolution of the cumulative wealth and its empirical volatility over time thorough a non parametric Markov bivariate process. As in Ortobelli Lozza, Angelelli, and Bianchi (2011) and Bean, Kontoleon, and Taylor (2008), a tree structure is used to represent the investment evolution under the Markovian hypothesis, but differently by their analysis here it is used a non parametric Markov chain process to approximate the portfolio return distribution and its volatility under a real world probability. Moreover, the *recombining effect* significantly reduce the problem complexity. This tree structure allows an appropriate analysis of the portfolio return and its empirical volatility joints tails. In this framework, the impact of the tails on the optimal portfolio choices is analyzed. Indeed, it is well known in the financial literature (see, among others, Rachev and Mittnik (2000) and the references therein) that the observed heavy tails of the return distributions could have a strong impact on the future wealth, since the return tails determine the probabilities of future losses and gains. Thus, the main contribution of this work is twofold. Firstly, a non parametric Markov stochastic volatility process is introduced and it is applied to portfolio selection problems. Secondly, the impact of the joint tails of the stochastic volatility model on optimal portfolio choices is examined through an empirical analysis.

In such analysis two alternative portfolio performance measures are considered: a Dynamic Sharpe Ratio, and a stochastic dynamic benchmark that takes into account the behaviour of a stochastic market benchmark.
In financial literature there exist several performance measures which are used either to measure the ex-post performance of portfolio strategies or to choose optimal portfolios in line with Sharpe thinking (see, among others, Sharpe (1994), Cogneau and Hübner (2009a) Cogneau and Hübner (2009b). Indeed, if the assumption of normality in return distributions is omitted, the classical risk?-reward Sharpe Ratio becomes a questionable tool for ranking risky projects. A general risk-?reward ratio suitable to compare skewed returns with respect to a stochastic benchmark should account asymmetric preferences to bet on potential high stakes and the aversion against possible huge volatility. The former goal is achieved by the proposed modified Sharpe ratio, where, as risk measure the expected portfolio volatility is adopted and, as reward measure the expected excess return (respect to the risk-less as for the classic Sharpe ratio) is used and it is conditioned to the forecasted portfolio wealth. It must be greater than the one obtained with the stochastic benchmark. Doing so, the optimization is still referred at the reward for unity of risk. The over performance with respect to the stochastic market benchmark is emphasized. Then, four joint tails of the bivariate stochastic volatility process are characterized. The process is conditioned to belong to a given tail of its bivariate distribution. The performance measures are then evaluated. Finally, the ex-post wealth obtained optimizing the conditional process is compared with respect to the unconditional portfolio performance ratios. Moreover, the presence of proportional transaction costs is assumed, this in order to construct a realistic investment strategy , as in Fu et al. (2015) and in Valladão, Silva, and Poggi (2018).

The remaining part of this chapter is organized as follows: in Section 2 describes the adopted methodology and it explain how it is built and computed the new stochastic volatility portfolio model. In Section 3 some portfolio strategies are examined and it is described the contribution of joint tail distribution of the proposed stochastic volatility model.

Section 4 empirically evaluates the impact on optimal choices of the bivariate Markov stochastic volatility tails.

1.2 Approximating a stochastic volatility model with a bivariate Markov chain

In this section the methodology adopted to construct the homogeneous Markov stochastic volatility portfolio process is described. It is used to derive the joint evolution of the cumulated wealth obtained investing in a portfolio up to time T and its cumulated stochastic volatility.

Let $Z_{x,t} = \sum_{i=1}^{n} x_i \frac{P_{i,t}}{P_{i,t-1}}$ be the portfolio of gross returns at time t, whereas the vector $x = [x_1, \ldots, x_n]'$ indicates the percentages of the initial wealth invested in each of the n assets and $P_t = [P_{1,t}, \ldots, P_{n,t}]'$ is the vector of prices at time t. No short sales are allowed. Thus, the vector x of portfolio weights belongs to the $(n-1)$-dimensional simplex $S = \{x \in \mathbb{R}^n \mid \sum_{i=1}^{n} x_i = 1; x_i \geq 0\}$.

For a portfolio $Z_{x,t}$, the empirical volatility is

$$\sigma_{Z_{x,t}} = \left(\sum_{i=t-m+1}^{t} \frac{\left(Z_{x,t} - \bar{Z}_{x,t}\right)^2}{m} \right)^{0.5}$$

where $\bar{Z}_{x,t}$ is the empirical mean computed on a time window of length m. For computational convenience the process of the empirical exponential volatility is modeled as: $Z_{\sigma,t} = \exp(\sigma_{Z_{x,t}})$. For any portfolio of weights x the stochastic volatility portfolio process $Z_t = (Z_{x,t}, Z_{\sigma,t})$ is an adapted bivariate Markov process defined on a filtered probability space $\left(\Omega, \Im, (\Im_t)_{0 \leq t \leq \infty}, \Pr\right)$, that it can be approximated with an homogeneous Markov chain. Clearly, the use of a bivariate process allows to better distinguish the proper contribution of portfolio returns and its stochastic volatility (and also of their joint distributional tails) in the optimal portfolio selection analysis . To build the Markov chain N states are considered for the gross returns portfolio and M states for its exponential volatility. Then, it is introduced the multi-index $i = (i_x, i_\sigma)$ to denote the states of the Markov chain $z^{(i)} = (z_x^{(i_x)}, z_\sigma^{(i_\sigma)})'$, $i \in I := \{(i_x, i_\sigma) : 1 \leq i_x \leq N, 1 \leq i_\sigma \leq M\}$. In particular, the states are defined discretizing the support of the Markov process $\{Z_t\}_{t \geq 0}$. Thus, given a set of past observations $\{Z_{-L}, \ldots, Z_0\}$, the range of the portfolio process is considered

$$\left(\min_{s=-L,\ldots,0} Z_{x,s}, \max_{s=-L,\ldots,0} Z_{x,s} \right) \times \left(\min_{s=-L,\ldots,0} Z_{\sigma,s}, \max_{s=-L,\ldots,0} Z_{\sigma,s} \right)$$

and it is divided into $N \cdot M$ bi-dimensional intervals $(a_i, a_{i-1}) \times (b_j, b_{j-1})$, where $\{a_i\}$ and $\{b_j\}$ are given by the two following decreasing sequences

$$\begin{aligned} a_i &:= u_x^i \max_s Z_{x,s}, \quad i = 0, \ldots, N \\ b_j &:= u_\sigma^j \max_s Z_{\sigma,s}, \quad j = 0, \ldots, M \end{aligned} \tag{6}$$

and $u_x := \left(\frac{\min_s Z_{x,s}}{\max_s Z_{x,s}}\right)^{1/N}$, $u_\sigma := \left(\frac{\min_s Z_{\sigma,s}}{\max_s Z_{\sigma,s}}\right)^{1/M}$ are two step factors useful to determine the process states.

Then, the values of the Markov process are approximated by the belonging to the bi-dimensional interval $(a_{i_x}, a_{i_x-1}) \times (b_{i_\sigma}, b_{i_\sigma-1})$ by the state $z^{(i)} = (z_x^{(i_x)}, z_\sigma^{(i_\sigma)})$ of the Markov chain defined by

$$\begin{aligned} z_x^{(i_x)} &= \sqrt{a_{i_x} a_{i_x-1}} = u_x^{\frac{1-2i_x}{2}} \max_k Z_{x,k}, \quad i_x = 1, \ldots, N \\ z_\sigma^{(i_\sigma)} &= \sqrt{b_{i_\sigma} b_{i_\sigma-1}} = u_\sigma^{\frac{1-2i_\sigma}{2}} \max_k Z_{\sigma,k}, \quad i_\sigma = 1, \ldots, M. \end{aligned} \tag{7}$$

Thus, according to (6) and (7), $z_x^{(i_x)} = z_x^{(1)} u_x^{1-i_x}$ and $z_\sigma^{(i_\sigma)} = z_\sigma^{(1)} u_\sigma^{1-i_\sigma}$. Clearly, for M states of the exponential volatility there are exactly M states of the portfolio volatility given by $\sigma_{(i_\sigma)} = \ln(z_\sigma^{(i_\sigma)})$, $i_\sigma = 1, \ldots, M$. The joint probability

$$\pi_i = \Pr\left(Z_{x,t} = z_x^{(i_x)}, Z_{\sigma,t} = z_\sigma^{(i_\sigma)}\right)$$

of the portfolio return and its exponential volatility respectively in states $z_x^{(i_x)}$ and $z_\sigma^{(i_\sigma)}$ is approximated as the number of times the process Z_t is in the state $z^{(i)} = (z_x^{(i_x)}, z_\sigma^{(i_\sigma)})$ divided by the total number of joint observations. As a consequence of the homogeneous property of the Markov chain the transition matrix is constant over the time and is given by $\Pi = [\pi_{ij}]_{i,j \in I}$, where:

$$\pi_{ij} = \Pr\left(Z_{\tau+1} = z^{(j)} | Z_\tau = z^{(i)}\right), \; i, j \in I$$

is the probability to move from the generic state $z^{(i)}$ to $z^{(j)}$ in one period of time. The estimates $\hat{\pi}_{ij}$ of these probabilities are obtained as the ratio of the number of observations that transit from state $z^{(i)}$ to state $z^{(j)}$ and the number of observations which are in the state $z^{(i)}$.

The proposed model is suitable to describe the joint evolution of the cumulated wealth obtained investing in a portfolio and its stochastic volatility. The bivariate cumulative process generated by the portfolio return is considered and its volatility. At time zero (no investment), the initial wealth is equal to 1 (i.e., $W_{0,x} = \sum_{i=1}^n x_i = 1$) and also its volatility is equal to 0 (i.e., $W_{0,\sigma} = 1 = \exp(0)$). The cumulative process $W_t = (W_{t,x}, W_{t,\sigma})'$ at time t is a bivariate random variable that can assume $N \cdot M$ possible values for any realized value of the cumulative process W_{t-1} at time $t-1$, i.e.,

$$W_t = z^{(i)} \otimes W_{t-1} = \left(z_x^{(i_x)} W_{(t-1),x}, \; z_\sigma^{(i_\sigma)} W_{(t-1),\sigma}\right)', \quad i = (i_x, i_\sigma) \in I.$$

In particular, $W_{t,x}$ and $W_{t,\sigma}$ point out the cumulative wealth and exponential volatility obtained investing in the portfolio x during the period $[0, t]$. Denoting $i(s) = (i_x(s), i_\sigma(s))$ the realized state of the Markov chain at time s, then a sample path of the cumulative value of the portfolio is described and its exponential volatility at time t can be seen as a function of the realized states $z^{(i(\tau))}$ in different times $\tau = 1, ..., t$, i.e.:

$$\overline{W}_t = \begin{pmatrix} W_{0,x} z_x^{(i_x(1))} z_x^{(i_x(2))} ... z_x^{(i_x(t))} \\ W_{0,\sigma} z_\sigma^{(i_\sigma(1))} z_\sigma^{(i_\sigma(2))} ... z_\sigma^{(i_\sigma(t))} \end{pmatrix}$$

where the cumulative process at time 0 is $W_0 = (W_{0,x}, W_{0,\sigma}) = (1, 1)$. Observe that, the largest and the smallest nodes of the wealth (and respectively of the volatility) grows and decreases exponentially and thus a much larger domain is covered for the future wealth and volatility taking also into account rare events. The sequence $\langle i(1), \ldots, i(t) \rangle$ identifies uniquely the path followed by the bivariate cumulative process up to time t. Moreover, the cumulative volatility during the period $[0, t]$ along the sample path, is given by the logarithm of the exponential volatility, i.e.

$$\ln(\overline{W}_{t,\sigma}) = \sum_{\tau=1}^{t} \ln\left(z_\sigma^{(i_\sigma(\tau))}\right) = \sum_{\tau=1}^{t} \sigma_{(i_\sigma(\tau))}.$$

Thus, considering all the possible sample paths, the average volatility process can be defined as:

$$\bar{\sigma}_{0,x} = 0, \text{ for } t = 0 \text{ and } \bar{\sigma}_{t,x} = \frac{\ln(W_{t,\sigma})}{t} \text{ for } t > 1. \tag{8}$$

Observe that the random variable $\bar{\sigma}_{t,x}$ (for any time $t > 1$) represents the forecasted average volatility over the period $[0, t]$. In the following portfolio selection problems the process $\{(W_{t,x}, \bar{\sigma}_{t,x})\}_{t \geq 0}$ is used and it is composed by the cumulative portfolio wealth $W_{t,x}$ and its average volatility process $\bar{\sigma}_{t,x}$.

A general bivariate Markov chain with $N \cdot M$ possible states should imply that the number of possible values for the cumulative process W_t grows exponentially with the time. However, as a consequence of the proposed construction, the process W_t can take only $[1 + t(N-1)] \cdot [1 + t(M-1)]$ values and the global number of the possible values of W_t up to time T is $\sum_{t=1}^{T} [1 + (N-1)t][1 + (M-1)t] = O(NMT^3)$. This property is called *recombining effect* of the Markov chain on the cumulative process W and it contributes to reduce the complexity of the problem.

This allows to describe the possible values of the cumulated wealth along the time dimension, using a tree structure according to 3 since, by (6) and (7), the possible values of W_t at time t can be denoted by:

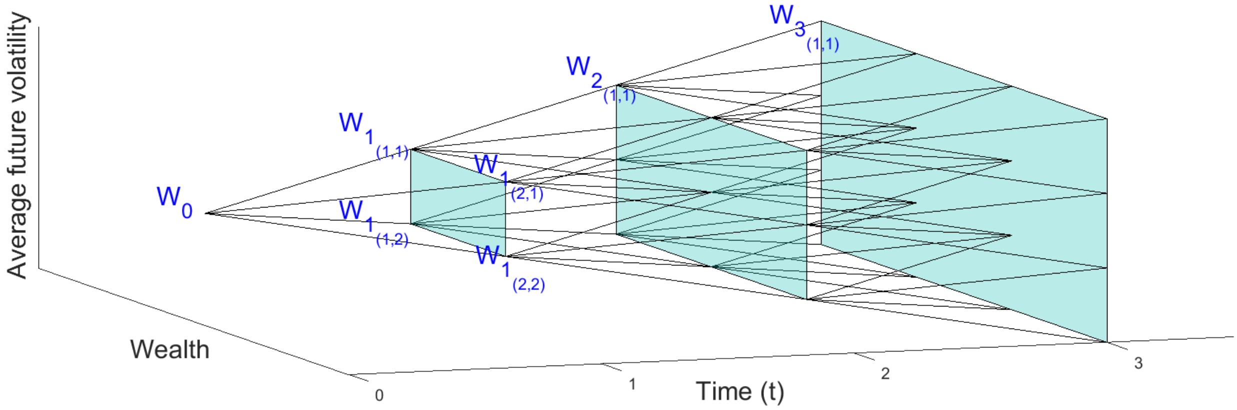

Figure 3: Graphical interpretation of the Markov tree

.

$$W_{t_{(l_x,l_\sigma)}} = \begin{pmatrix} W_{t,x,l_x} \\ W_{t,\sigma,l_\sigma} \end{pmatrix} = \begin{pmatrix} (z_x^{(1)})^t u_x^{1-l_x} \\ (z_\sigma^{(1)})^t u_\sigma^{1-l_\sigma} \end{pmatrix} \tag{9}$$

where

$$(l_x, l_\sigma) \in A_t := \{(l_x, l_\sigma) : 1 \le l_x \le 1 + t(N-1), 1 \le l_\sigma \le 1 + t(M-1)\},$$

that is the values W_{t,x,l_x} and W_{t,σ,l_σ} are decreasing functions respectively of l_x and l_σ. A bivariate tree can intuitively be used to describe the evolution of the cumulative process W_t, starting with a single node $W_{0_{(1,1)}} = (1,1)'$ at time 0 and presenting at each time instant t the $[1+t(N-1)] \times [1+t(M-1)]$ nodes given by W_{t_l}, $l = (l_x, l_\sigma) \in A_t$. In 3 it is shown how the recombining effect acts on the process when a bivariate tree with $N = M = 2$ is reproduced.
Each colored polygon in the tree represents the regions of the possible realizations of our discretised process, where at each time $(t = 0, ..., 3)$ we have $[1+t]^2$ nodes of the cumulative process W_t.
The joint probability of obtaining the cumulative values W_{t_l} (for any $l \in A_t$) in state $z^{(i)}$ (for any $i \in I$) at time t is computed as:

$$\pi_{(W_t,Z_t)}(l,i) = \Pr(W_t = W_{t_l} \cap Z_t = z^{(i)}).$$

These probabilities can be computed recursively by the formula
$\pi_{(W_t,Z_t)}(l,i) =$

$$\begin{cases} \pi_i & t = 0, l = \mathbf{1} = (1,1) \\ \sum_{h \in I} \pi_{(W_{t-1},Z_{t-1})}(l - (i - \mathbf{1}), h)\pi_{hi} & t > 0, l_x - (i_x - 1) > 0, \\ & \text{and } l_y - (i_y - 1) > 0 \\ 0 & otherwise. \end{cases}$$

Then, the probability $\pi_{W_t}(l)$ is simply given by

$$\pi_{W_t}(l) = \begin{cases} 1 & t = 0, l = (1,1) \\ \sum_{h \in I} \pi_{(W_t, Z_t)}(l, h) & t > 0 \\ 0 & otherwise \end{cases} \tag{10}$$

These probabilities characterize both the distributions of the cumulative process W_t and of the process $\{(W_{t,x}, \bar{\sigma}_{t,x})\}_{t \geq 0}$. Moreover, the proposed procedure can also be used to approximate the joint markovian behaviour of one portfolio and a given benchmark $Z_{(x,b),t} = (Z_{x,t}, Z_{b,t})$. This leads to a cumulative wealth process $W_{(x,b)} = \{(W_{t,x}, W_{t,b})\}_{t \geq 0}$ as described by Ortobelli Lozza, Angelelli, and Bianchi (2011).

1.3 Joint tails in portfolio selection stochastic volatility models

This section propose different strategies for portfolio selection when the portfolio of gross returns and its volatility is approximated by a bivariate Markov chain as in Section 2. In particular, firstly some portfolio performance measures are examined, then it is discussed how these performance measures can be implemented in order to take into account the joint tails of the cumulative wealth process and its average volatility process.

In portfolio theory, typically it is maximized a functional g (performance measure or utility functional), that depends on the portfolio weights x belonging to the $(n-1)$-dimensional simplex S (when no short sales are allowed). Clearly, the proper choice of the functional g is related to the investor's preferences and, for this reason, it is required that g is isotonic with the preferences of a particular class of investors (i.e., $g(X) > g(Y)$ any time X is preferred to Y by a given class of investors) see for details Angelelli and Ortobelli Lozza (2009). A static portfolio problem is any optimization problem that do not consider the time evolution of the wealth and for which the utility functional g is applied to portfolio of gross returns $Z_{x,t}$. In these cases the solution of the problem is:

$$\max_{x \in S} g(Z_{x,t}).$$

Probably the most known and used performance functional g is the *Sharpe Ratio (SR)* Sharpe (1994) which evaluates the expected excess return for unit of risk, measured as the standard deviation, i.e.

$$SR(Z_{x,t}) = \frac{E(Z_{x,t} - r_{rf})}{\sigma_{Z_{x,t}}},$$

where r_{rf} is the risk free return and $\sigma_{Z_{x,t}}$ is the standard deviation of the portfolio of gross return. SR can be seen as a risk-to variability measure.

The Sharpe ratio is isotonic with non-satiable risk averse preferences (i.e., any time X is preferred to Y by all non-satiable risk averse investors, then $SR(X) > SR(Y)$).

A dynamic portfolio problem generally optimize a functional over S taking into account the time evolution of portfolio wealth. In particular, under the assumption of Section 2, the trade-off between the forecasted portfolio wealth and its average volatility at a given temporal horizon T can be optimized. In this application several portfolio selection strategies are considered. Investors maximize a functional $g(\cdot)$ every T periods applied to the cumulative process W_T (or to the sample path of the cumulative process) evaluated at time T. Thus investors periodically compute the optimal portfolio weights $x_M \in S$ such that

$$x_M = arg \max_{x \in S} g(W_T).$$

Initially it is considered a dynamic version of the Sharpe ratio given by:

$$DSR(W_T) = E\left(\frac{W_{T,x} - W_{T,rf}}{\bar{\sigma}_{T,x}}\right), \tag{11}$$

where $W_{T,rf}$ is the final wealth obtained investing in the risk free asset and $\bar{\sigma}_{T,x}$ is the average volatility over the period $[0, T]$ defined in (8).

In general, for a bivariate cumulative process $\{W_t\}_{t \geq 0}$ the joint dynamic of the portfolio cumulative wealth is considered with its average volatility, where $W_t = (W_{t,x}, W_{t,\sigma})$ and so

$$E(f(W_t)) = \sum_{l \in A_t} f(W_{t_l}) \pi_{W_t}(l)$$

where W_{t_l}, $l \in A_t$ are the $[1 + t(N-1)] \times [1 + t(M-1)]$ nodes.
Similarly, the conditional expected value is

$$E(f(W_t)\,|W_t \in C\,) = \frac{\sum_{l \in C} f(W_{t_l}) \pi_{W_t}(l)}{\sum_{l \in C} \pi_{W_t}(l)}. \tag{12}$$

Thus, according to (9) and (10), considering the portfolio cumulative wealth and its average volatility , the above dynamic Sharpe ratio can be evaluated as

$$DSR(W_T) = \sum_{l_x=1}^{1+T(N-1)} \sum_{l_\sigma=1}^{1+T(M-1)} \frac{W_{T,x,l_x} - W_{T,rf}}{\bar{\sigma}_{T,x,l_\sigma}} \pi_{W_T}(l_x, l_\sigma),$$

where $\bar{\sigma}_{T,x,l_\sigma} = \frac{\ln\left(W_{T,\sigma,l_\sigma}\right)}{T}$ is the average volatility over the period $[0, T]$ obtained by the node W_{T,σ,l_σ} of the cumulative exponential volatility.

In portfolio theory, since the Sharpe's statements, several alternative performance measures have been proposed (see Cogneau and Hübner (2009a)Cogneau and Hübner (2009b)) in order to account for investors' preferences. In particular, when investors consider a stochastic benchmark characterised by a final wealth $W_{T,b}$, they optimize the ratio of expected gains with respect to the benchmark, over the risk. In this context the Sharpe ratio can be modified taking into account only the forecasted wealth greater than the stochastic benchmark.
This is a reward–to–variability ratio, and it represents the potential for positive returns compared to the volatility. Our performance measure is in principle very similar to the Sharpe ratio except that for the market stochastic benchmark conditioning on the excess return. It is proposed the use the following alternative stochastic benchmark ratio (in short SBR).

$$SBR(W_T) = \frac{E\left(W_{T,x} - W_{T,rf}\,|W_{T,x} \geq W_{T,b}\right)}{E\left(\bar{\sigma}_{T,x}\right)}, \tag{13}$$

The stochastic benchmark ratio evaluates the expected excess positive wealth using the cumulative wealth process $W_t = \{(W_{t,x}, W_{t,b})\}_{t\geq 0}$ to account for the joint behaviour of the portfolio and the benchmark over the portfolio stochastic volatility. Thus, the conditional expected value of the stochastic benchmark ratio (13) can be computed as

$$E\left(W_{T,x} - W_{T,rf}\,|W_{T,x} \geq W_{T,b}\right) = \frac{\sum_{(l_x,l_b)\in C}\left(W_{T,x,l_x} - W_{T,rf}\right)\pi_{W_T}(l_x, l_b)}{\sum_{(l_x,l_b)\in C}\pi_{W_T}(l_x, l_b)}$$

where the conditioning region is given by $C = \{(l_x, l_b)\,|W_{T,x,l_x} \geq W_{T,b,l_b}\}$ and according to (9) and (10), the above stochastic performance ratio is given by

$$SBR(W_T) = \frac{\sum_{(l_x,l_b)\in C}\left(W_{T,x,l_x} - W_{T,rf}\right)\pi_{W_T}(l_x, l_b)}{\sum_{(l_x,l_b)\in C}\pi_{W_T}(l_x, l_b)\sum_{l_\sigma=1}^{1+T(M-1)}\bar{\sigma}_{T,x,l_\sigma}\sum_{l_x=1}^{1+T(N-1)}\pi_{W_T}(l_x, l_\sigma)}, \tag{14}$$

where:

1. $\bar{\sigma}_{T,x,l_\sigma} = \frac{\ln\left(W_{T,\sigma,l_\sigma}\right)}{T}$ is the average volatility over the period $[0, T]$
2. $C = \{(l_x, l_b)\,|W_{T,x,l_x} \geq W_{T,b,l_b}\}$.

1.3.1 Tail portfolio performance measures

Generally financial models take into account the distributional tails, conditioning the stochastic variables to belong to the tails. Thus, the impact of the portfolio and volatility joint tails on the optimal portfolio choices can be evaluated. In practice, starting by the previous performance measures (11) and (13), two alternative performance ratios can be obtained: the conditional dynamic Sharpe ratio (namely, C_DSR) and the conditional stochastic benchmark ratio (namely, C_SBR) given by:

$$C\_DSR(W_{T,B}) = E\left(\frac{W_{T,x} - W_{T,rf}}{\bar{\sigma}_{T,x}} \middle| (W_{T,x}, \bar{\sigma}_{T,x}) \in B \right) \tag{15}$$

$$C\_SBR(W_{T,(x,b),B}) = \frac{E\left(W_{T,x} - W_{T,rf}\,|W_{T,x} \geq W_{T,b}\right)}{E\left(\bar{\sigma}_{T,x-}|\,(W_{T,x}, \bar{\sigma}_{T,x}) \in B\right)} \tag{16}$$

where the set B points out a tail area of the joint stochastic volatility process $(W_{t,x}, \bar{\sigma}_{t,x})$. In particular, are defined four admissible tail areas whose joint probability is fixed equal to β, (i.e. $\Pr\left((W_{T,x}, \bar{\sigma}_{T,x}) \in B\right) = \beta$) which are:

1. High wealth and low volatility (namely, H-L zone), where $B = (W_{T,x} > c; \bar{\sigma}_{T,x} \leq d)$ and the values c, and d are identified such that $\Pr(W_{T,x} > c) = \Pr(\bar{\sigma}_{T,x} \leq d)$;
2. High wealth and high volatility (namely, H-H zone), where $B = (W_{T,x} > c; \bar{\sigma}_{T,x} > d)$ and the values c, and d are identified such that $\Pr(W_{T,x} > c) = \Pr(\bar{\sigma}_{T,x} > d)$;
3. Low wealth and high volatility (namely, L-H zone), where $B = (W_{T,x} \leq c; \bar{\sigma}_{T,x} > d)$ and the values c, and d are identified such that $\Pr(W_{T,x} \leq c) = \Pr(\bar{\sigma}_{T,x} > d)$;
4. Low wealth and low volatility (namely, L-L zone), where $B = (W_{T,x} \leq c; \bar{\sigma}_{T,x} \leq d)$ and the values c, and d are identified such that $\Pr(W_{T,x} \leq c) = \Pr(\bar{\sigma}_{T,x} \leq d)$.

Therefore, if for example the dynamic Sharpe ratio is computed conditioned on the H-L zone the result is

$$C\_DSR(W_{T,B}) = \frac{\sum_{(l_x,l_\sigma)\in C_{H-L}} \frac{W_{T,x,l_x} - W_{T,rf}}{\bar{\sigma}_{T,x,l_\sigma}} \pi_{W_T}(l_x, l_\sigma)}{\sum_{(l_x,l_\sigma)\in C_{H-L}} \pi_{W_T}(l_x, l_\sigma)}$$

where the region C_{H-L} is the set of nodes (l_x, l_σ) such that

$$\sum_{l_x=k_c}^{1+T(N-1)} \sum_{l_\sigma=1}^{k_d} \pi_{W_T}(l_x, l_\sigma) = \beta;$$

and k_c, k_d are determined such that

$$\sum_{l_\sigma=1}^{k_d} \sum_{l_x=1}^{1+T(N-1)} \pi_{W_T}(l_x, l_\sigma) = \sum_{l_x=k_c}^{1+T(N-1)} \sum_{l_\sigma=1}^{1+T(M-1)} \pi_{W_T}(l_x, l_\sigma).$$

Clearly, the equalities in probability introduced above, to determine the tail areas, can be satisfied if the probability distributions are continuous. This is not this case since the discretization of the distributions has been done with the procedure described in Section 2.
Thus, the values c and d are nodes properly chosen to approximate the above equalities. For example, to consider the H-L zone $c = W_{T,x,l_x}$ and $d = \bar{\sigma}_{T,x,l_\sigma}$ must be chosen such that one of the following conditions hold:

a)

$$\Pr(W_{T,x} > W_{T,x,l_x}; \bar{\sigma}_{T,x} \leq \bar{\sigma}_{T,x,l_\sigma}) \geq \beta,$$

and

$$\Pr(W_{T,x} > W_{T,x,l_x-1}; \bar{\sigma}_{T,x} \leq \bar{\sigma}_{T,x,l_\sigma+1}) < \beta;$$

b)

$$\Pr(W_{T,x} > W_{T,x,l_x}) \geq \Pr(\bar{\sigma}_{T,x} \leq \bar{\sigma}_{T,x,l_\sigma})$$

and

$$\Pr(W_{T,x} > W_{T,x,l_x-1}) < \Pr(\bar{\sigma}_{T,x} \leq \bar{\sigma}_{T,x,l_\sigma}),$$

or

$$\Pr(\bar{\sigma}_{T,x} \leq \bar{\sigma}_{T,x,l_\sigma}) \geq \Pr(W_{T,x} > W_{T,x,l_x})$$

and

$$\Pr(\bar{\sigma}_{T,x} \leq \bar{\sigma}_{T,x,l_\sigma+1}) < \Pr(W_{T,x} > W_{T,x,l_x}).$$

Proceeding in a similar way it is possible to determine the values c and d for all the other tail zones.

1.4 An empirical analysis of the stochastic volatility portfolio model

In this section it is investigated the impact of the tails in portfolio selection via the bivariate Markov stochastic volatility model. In particular, it is presented the comparison among some portfolio strategies based on the maximization either of the static Sharpe ratio or of the proposed portfolio performance measures (DSR and SBR) that account for the portfolio behavior on the distribution tails. First, it is proposed the out of sample valuation of the wealth produced optimizing the Dynamic Sharpe ratio and the Stochastic benchmark ratio. Second, it is examined the distributional behavior of optimal portfolios conditioned on the different distributional tails. In order to perform this analysis it is used the set of adjusted closing prices of the Standard and Poor 500 Index (our benchmark) and the components of DJIA index from 12th October 1998 to 22nd May 2019. Moreover, the 3M Treasury Bill is assumed as risk–less asset. It is set $N = M = 4$ states for each asset and $\beta = 1\%$. The Markov model is calibrated on a one-year time window of historical daily observations (252 trading days). For each window, the first 6-months is used to compute the initial empirical standard deviation and the last 6-months (126 trading days) to calibrate the Markov model. The optimal portfolio is selected on a time horizon of 21 days (i.e. $T = 21$, one month). Finally window is translated by 21 days and the portfolio is monthly re-calibrate for a total of 234 calibrations.

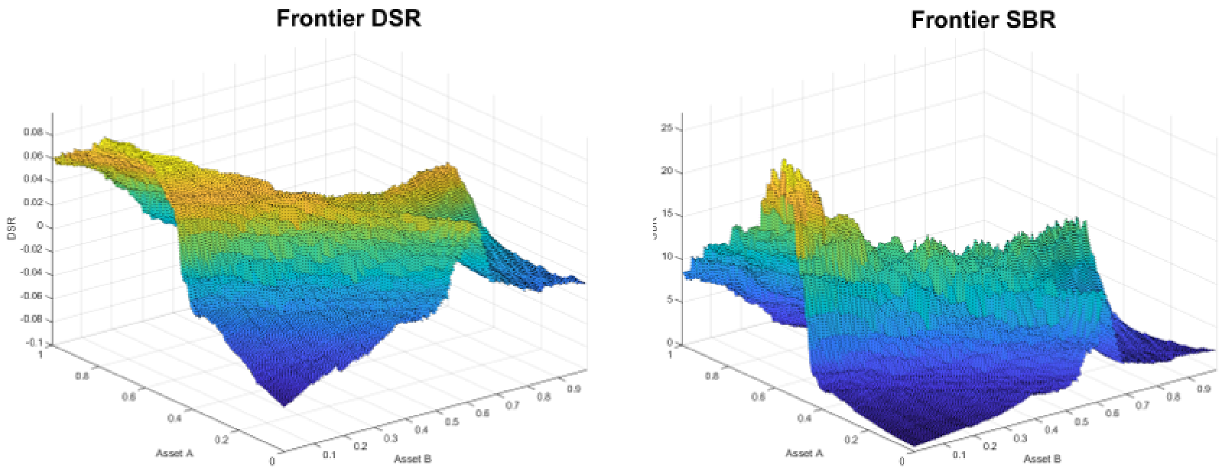

Figure 4: Dynamic Sharpe Ratio (DSR) and Stochastic Benchmark ratio (SBR) computed on a portfolio of three assets

Just in order to have a graphical interpretation, in figure 4 it is reported the value of the Dynamic Sharpe Ratio (DSR) and of the Stochastic Benchmark ratio (SBR) computed on a toy-portfolio composed by only three assets (Microsoft, Coca Cola, Boeing) changing the weights belonging to $S = \left\{x \in \mathbb{R}^3 \mid \sum_{i=1}^{3} x_i = 1; x_i \geq 0\right\}$. The performance measures based on the Markov hypothesis provide several local optima. Therefore, the search for an optimal solution is often faced with problems related to the computational complexity of the maximization of these performance measures.

For this reason, in these experiments it is first applied an optimizer based on the idea presented in Angelelli and Ortobelli Lozza (2009) to find a proper starting point. Then, the different portfolio problems are optimized using the algorithm "pattern search" presented in Matlab libraries. A comparison analysis of the combined use of these two optimizers allows to save time and get better results than other global optimizers essentially based or on Genetic algorithms or on the Simulated Annealing type algorithms.

1.4.1 Portfolio selection comparison

In the first empirical analysis it is considered an out of sample comparison among the different strategies which correspond to the different performance measures. In particular, for each strategy, it is set an initial wealth $W_0 = 1$ and it is assumed that no short sales are allowed. Thus, starting from October 12th 1999 the optimal portfolio $x^{(K)}$, $k = 0, 1, 2, ..., 234$ it is determined by recalibrating every month (21 trading days). At each calibration k, two main steps are repeated for all the performance measures, in order to compute the out of sample final wealth obtained by the different strategies:

- **Step 1** Determine the optimal portfolio $x^{(k)}$ that maximizes the performance ratio $\rho(W(x))$ (SR, DSR, SBR, C_DSR, C_SBR) associated to the relative strategy, i.e. the solution of the following optimization problem:

$$\begin{array}{l} \max\limits_{x^{(k)}} \rho(W_T(x^{(k)})) \\ \text{s.t.} \\ \sum_{i=1}^{n} x_i^{(k)} = 1, \\ x_i^{(k)} \geq 0; \quad i = 1, \ldots, n. \end{array}$$

  In more details, the initial portfolio weights vector is determined using the heuristic by Angelelli and Ortobelli Lozza (2009). Then, the optimal solution is obtained by a pattern search algorithm: at each iteration, the 2n directions given by the canonical basis and its opposite are explored. Finally the closest point to the current solution which satisfies working set constraints and provide the best improvement of the objective function is selected as new current solution. The value of the objecting function $\rho(cdot)$ is computed on the forecasted wealth $W_T(x)$ obtained by the portfolio with composition (x).

- **Step 2** During the period $[t_k, t_{k+1}]$ (where $t_{k+1} = t_k + T$) the portfolio is re-calibrated considering 20 basis point of proportional transaction costs. Thus, the ex-post final wealth is given by:

$$W_{t_{k+1}} = W_{t_k} \left(\left(x^{(k)}\right)' z_{(t_{k+1})}^{(ex\ post)} - t.c. \left(x^{(k)}\right) \right), \tag{17}$$

  where:

  1. $z_{(t_{k+1})}^{(ex\ post)}$ is the vector of observed gross returns during the period from t_k to t_{k+1}
  2. $t.c.\left(x^{(k)}\right) =$

$$0.002 \sum_{i=1}^{30} \left| x_i^{(k)} - \frac{x_i^{(k-1)} z_{(t_k),i}}{\sum_{j=1}^{30} x_j^{(k-1)} z_{(t_k),i}} \right|$$

     are the proportional transaction costs.

The optimal portfolio $x^{(k)}$ is the new starting point for the $(k+1)$-th optimization problem and $W_{t_{k+1}}$ is the cumulative wealth to reinvest.

The results of the comparison of the different strategies are reported in figures 5, 6, 7 and table 1. Figure 5 reports the sample path of the out of sample wealth obtained with the static Sharpe ratio (1.3), the dynamic Sharpe ratio (11) and the *Standard* & *Poor* 500 index. In particular, figure 5 shows that the dynamic Sharpe ratio presents a better performance with respect to the index and to the static Sharpe ratio, since it is more capable to forecast the crisis period (i.e., the losses during the subprime crisis are smaller). However, remarkable differences between Sharpe type strategies can be found.

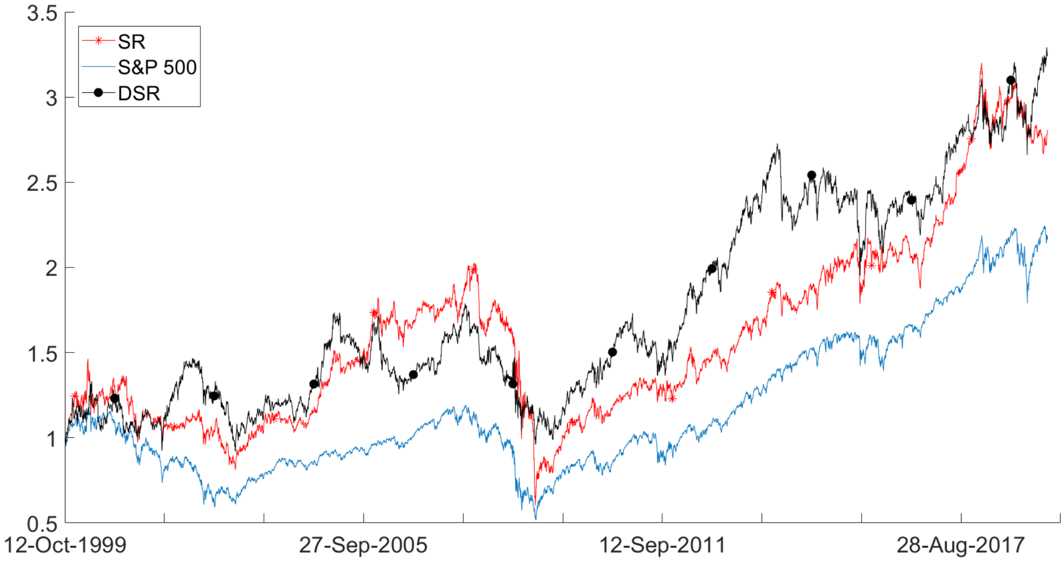

Figure 5: Comparison among cumulated wealth with Sharpe type performance strategies

Figures 6 and 7 report respectively the sample path of ex post wealth obtained with the conditional dynamic Sharpe ratio and the conditional stochastic benchmark ratio. The dynamic Sharpe ratio and all the conditional DSR strategies outperform the benchmark *Standard & Poor 500* and obtain an increment of wealth that is always greater than 3 times the initial one.

Similarly, figure 7 shows that the benchmark type strategies always outperform the benchmark. In particular, two strategies (namely in the H-L, H-H tails) obtain an increment of wealth of about 10 times the initial one while in the L-H and L-L tail, an increment of wealth is quantified in around 6 times the initial one but it is however lower than SBR performance (obtained without conditioning to a tail). Moreover, among the C_SBR portfolio strategies the best strategies until the crisis are the ones implemented in the areas with low volatility while, after this event,

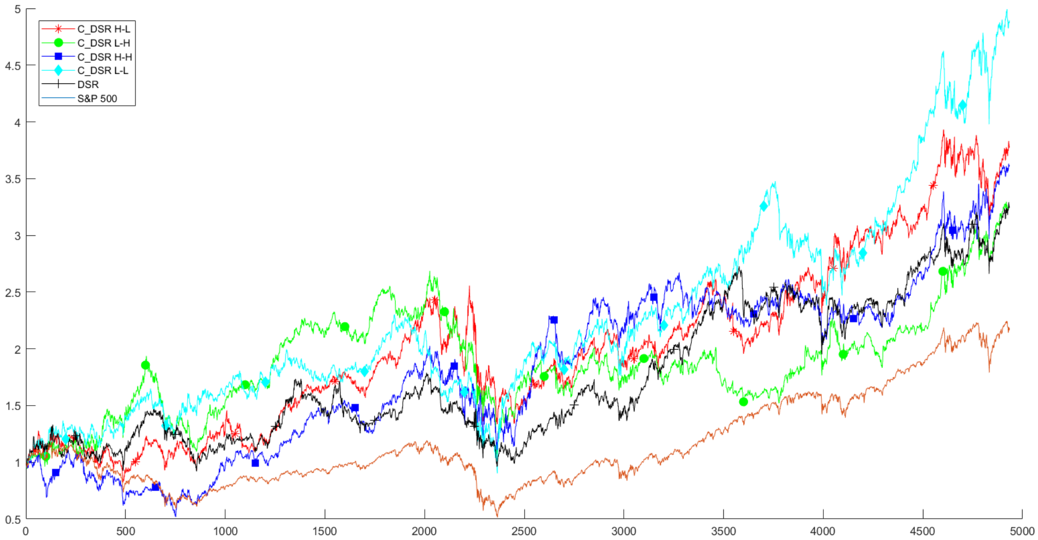

Figure 6: Comparison among conditional dynamic Sharpe type strategies

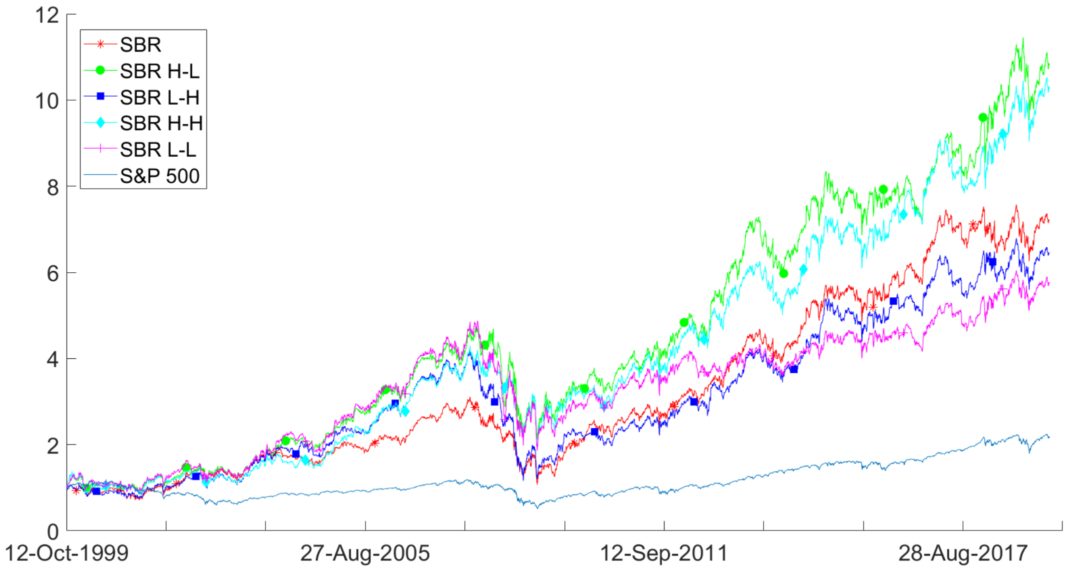

Figure 7: Comparison of conditional stochastic benchmark strategies

the C_SBR strategies on the high volatility tails performed better. C_SBR strategies perform better than the C_DSR ones even if it is observed a significative impact on the optimal choices for all conditional Sharpe type strategies.

The ordering among the results of this analysis, although evident from a graphical inspection of figures, have been tested to investigate the presence of some kind of stochastic dominance among the strategies. In particular, results are checked for the presence of stochastic dominance of the first two stochastic dominance orders (consistent with preferences of non-satiable investors, i.e. FSD, and non–satiable risk averse investors, i.e. SSD) and the Increasing-Convex-Order (consistent with preferences of non-satiable risk seeker investors, namely ICX). In Table 2.1 with the capital letters are described the strategies without the benchmark, with the lower case letters are described the strategies conditioned to the value of the Benchmark. It can be seen that there is not any first order stochastic dominance (FSD) and only in two cases it can be observed the second order stochastic dominance. On the one side, it appears evident that Conditional type strategies generally dominate the benchmark at least by the point of view of all non satiable risk seeker investors. On the other side, it is found that benchmark based strategies perform better than those based on the dynamic Sharpe ratio.

Table 1: Stochastic dominance among strategies and Benchmark.

	H–L	L–H	H–H	L–L	h–l	l–h	h–h	l–l	$S\&P500$
H–L	-	-	-	-	-	-	-	-	-
L–H	-	-	-	-	-	-	-	-	-
H–H	ICX	ICX	-	-	-	-	-	-	ICX
L–L	-	-	SSD	-	-	-	-	-	ICX
h–l	ICX	ICX	-	-	-	SSD	-	-	ICX
l–h	ICX	ICX	-	ICX	-	-	-	-	ICX
h–h	ICX	ICX	-	-	-	-	-	-	ICX
l–l	ICX	ICX	-	-	-	-	-	-	ICX
$S\&P500$	-	-	-	-	-	-	-	-	-

1.4.2 On the impact of the portfolios optimized on the tails

As observed in Section 2.4.1, conditioning the analysis on the tails an out of sample wealth higher than the corresponding unconditional can generally be observed. In this section it is clarified and justified the impact of the tails on the optimal choices. For illustrative purpose,the joint distribution of the final wealth and the average volatility of the optimal portfolios obtained optimizing the DSR in one year are graphically examined. Then, the behaviour of all optimal portfolios obtained is summarized in Section 2.4.1.

Figures 8 and 9 report the joint distributions of four optimal portfolios. They are get optimizing the DSR during the last year of observations, namely from May 23rd 2018 to May 22nd 2019. First of all, it is observed that the optimal choices are very different with respect to the others. Looking figures 8 and 9 it can be noted that the variability of the wealth appears very high, while the variability of the volatility seems very low (most of the forecasted volatility is concentrated on a unique zone).

For a more accurate analysis, in table 2 it is reported the ex-ante average descriptive statistics (mean, standard deviation, kurtosis and skewness) of the marginal distributions of the forecasted final wealth and average volatility obtained optimizing the conditional performance

Table 2: Average ex-ante statistics of the forecasted final wealth and volatility for the optimal portfolio of the conditional Sharpe type strategies

	C_DSR H–L	C_DSR L–H	C_DSR H–H	C_DSR L–L	C_SBR H–L	C_SBR L–H	C_SBR H–H	C_SBR L–L
W Stat								
μ	0.0103	0.0116	0.0094	0.0107	0.0103	0.0110	0.0100	0.0106
σ	0.0722	0.0716	0.0866	0.0685	0.0841	0.0904	0.0873	0.0854
Skew	0.1705	0.1550	0.2595	0.1474	0.1804	0.2368	0.1780	0.1864
Kurt	3.6295	3.7008	3.7910	3.6986	3.7507	3.7426	3.6692	3.7050
VaR5%	0.0947	0.0915	0.1040	0.0917	0.1112	0.1137	0.1151	0.1136
Volatility Stat								
μ	0.0129	0.0125	0.0141	0.0126	0.0143	0.0145	0.0144	0.0145
σ	0.0039	0.0038	0.0043	0.0037	0.0042	0.0044	0.0045	0.0043
Skew	0.4340	0.4127	0.3797	0.4071	0.4997	0.5448	0.5116	0.5076
Kurt	2.3655	2.2939	2.3355	2.4465	2.3972	2.4835	2.4913	2.4478
$\rho_{(W,var)}$	0.0777	0.0905	0.1006	0.0368	0.0513	0.0776	0.0439	0.0656

measures over all the examined 234 rebalancing periods. In addition come other statistics are considered:

- the value at risk ($VaR_p(X) = -F_X^{-1}(p) = -\inf\{s|\Pr(X \leq s) \geq p\}$,
- the conditional Value at risk ($\text{CVaR}_p(X) = \frac{1}{p}\int_0^p VaR_u(X)du$) of the final wealth $W_{T,x}$ both with a 95% confidence level (i.e., $p = 1 - 0.95$)
- the correlation between the wealth $W_{T,x}$ and average volatility $\bar{\sigma}_{T,x}$ (on average among all the optimal portfolios).

From these statistics can be argued that the C_DSR strategies generally present lower risk in terms of $\text{VaR}_{5\%}$ and $\text{CVaR}_{5\%}$ than C_SBR strategies. Moreover, the correlation between the forecasted wealth $W_{T,x}$ and its average volatility $\bar{\sigma}_{T,x}$ is generally very low, but it is higher for CSBR strategies.
For all strategies, and also according to figures 8 and 9, the standard deviation of the final wealth $W_{T,x}$ is almost 6 times greater than its average (i.e., high final wealth variability) and it is much more larger than the standard deviation of the volatility $\bar{\sigma}_{T,x}$.
Moreover, both the volatility $\bar{\sigma}_{T,x}$ and wealth $W_{T,x}$ present a positive skewness (that is larger for the volatility) and a leptokurtic final wealth is oserved on average (kurtosis of $W_{T,x}$ is on average greater than 3) and a platykurtic volatility (kurtosis of $\bar{\sigma}_{T,x}$ is on average smaller than 3). The C_DSR strategy valued on the L-H tail presents the lowest risk (among all strategies) in terms of standard deviation of final wealth, $\text{VaR}_{5\%}$, $\text{CVaR}_{5\%}$ and average of the forecasted variability.
This result is reasonable since with this strategy it is implicitly minimized the risk in a tail where the wealth is small and the volatility is high (i.e., in the worst case). According to this analysis the C_DSR strategy valued on the L-H tail presents the most risk conservative

situation probabily because optimize the risk on the worst wealth-risk situation, in a sort of MINMAX solution.

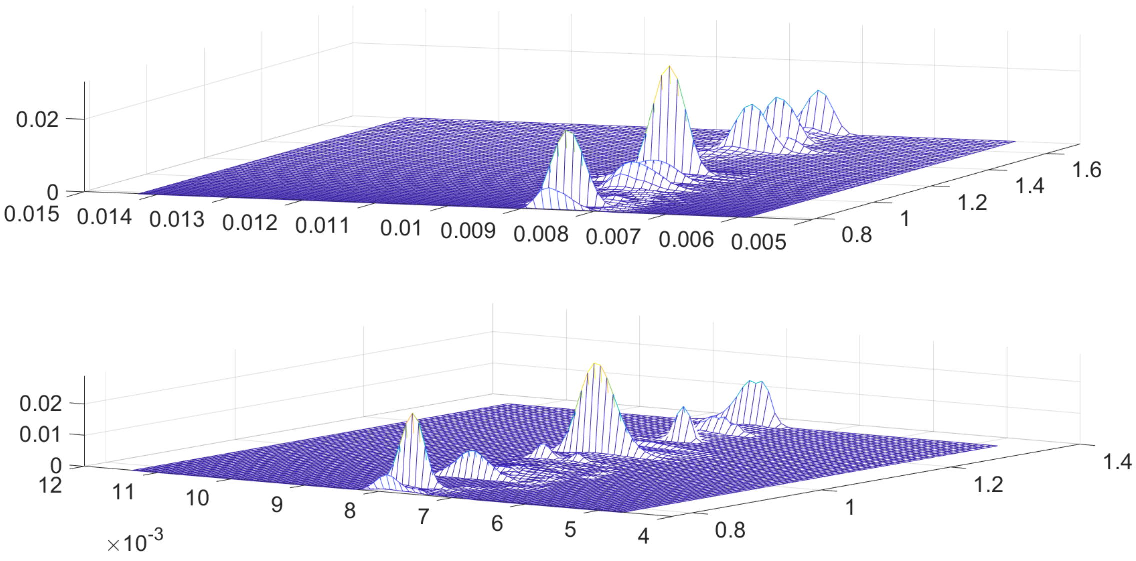

Figure 8: Joint distributions (wealth and variance at time T) of HL and HH optimal DSR portfolios

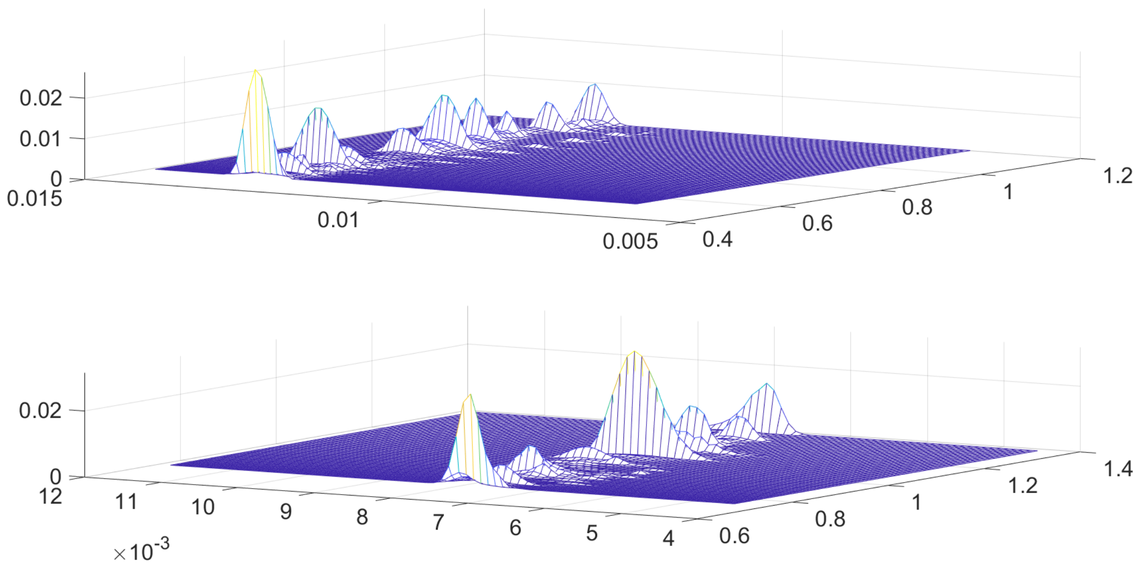

Figure 9: Joint distributions (wealth and variance at time T) of LL and LH optimal DSR portfolios

1.5 Conclusions

This chapter proposes a simple way to value stochastic volatility portfolio models using a bivariate Markov chain. In practice, with a discrete time model it is suggest how to examine the joint behavior of the future wealth and its average volatility. Moreover, alternative portfolio selection models are examined according to this modelization with and without taking into account the joint behavior of the wealth and its variability on one of their tails. An empirical comparison analysis show that taking into account the stochastic volatility is useful even under the optimization of the Sharpe ratio. Moreover conditioning the portfolio performance or the risk on one of the joint tails has generally a positive impact on the out of sample wealth. Finally, some ex-ante statistics on the optimal portfolios are discussed to examine and justify some of the ex post obtained results. This is just an initial point of the analysis of the problem but this proposed stochastic volatility model is promising and it can be also useful to deal other financial problems where it is required the joint modelization of the wealth and the volatility.

2. Enhanced tracking error quantile regression

2.1 Introduction

In 1970 was formulated the hypothesis of efficient markets (EMH) (see Malkiel and Fama, 1970) with the idea that, over long periods, financial indices are unbeatable, promoting the passive investments. In 2004 Lo (2004), in order to state the connection between EMH and behavioural finance, postulates that markets usually react rationally and instantaneously to new information, but in some situations they can be driven by "fear and greed". This allows the existence of asymmetries in markets and the possibility to outperform indices. According to these different views, portfolios can be constructed to passively replicate an index, while others are actively managed in order to generate active-returns without considering a precise benchmark. In between, there are semi passive strategies which mimic the benchmark looking for extra-performances, capturing the benefits of passive and active management. These strategies are indicated as enhanced indexing.

Recently, a branch of the financial literature addresses enhanced index strategies, proposing interesting models designed to obtain extra-performances in the index tracking framework. These strategies aim to outperform the index by generating "excess return" according to some other conditions, usually referred to the risk of portfolio.

According to Valle, Roman, and Mitra (2017), the return distribution of a replicating portfolio is considered enhanced if the left tail is improved, the downside risk is reduced and the standard deviation remains within a specified range.

In recent years, the process to build portfolios that mimic a given index looking for an extra performance is getting crucial. Investments into semi-passive, enhancing and systematic strategies are increasing. However, very few (see Mitra et al., 2018) enhanced indexation methods have been proposed (for an exhaustive survey of literature (see Canakgoz and Beasley, 2009)). Canakgoz and Beasley (2009) propose a regression based model for enhanced indexing, developing a two-stage mixed-integer linear programming approach in which they respectively focus on slope-intercept and transaction cost. In the first stage, they solve a problem achieving a regression slope as close as possible to one. This optimization is subjected to a constraint on the regression intercept. The second stage is focused on the minimization of transaction costs. In 2013, Roman, Mitra, and Zverovich (2013) apply a second order stochastic dominance strategy to construct a portfolio whose return distribution dominates the benchmark one. They adopt a multi-objective linear problem solved with a cutting-plane solution method presented in Fábián et al. (2011).

In 2011, Meade and Beasley (2011) investigate a momentum strategy via maximization of a modified Sortino ratio (see Sortino and Price, 1994) objective function. In 2014, Guastaroba et al. (2016) introduce a mixed-integer linear programming to enhance the index tracking problem maximizing the Omega ratio Keating and Shadwick (2002) in a linear formulation with buy-in threshold limits and cardinality constraints. Bruni et al. (2014) propose a linear bi-objective optimization approach to maximize the average excess return minimizing the risk. Afterwards they investigate the theoretical condition to guarantee the existence of an enhanced index portfolio.

Recently, with the aim to include behavioural finance in tracking models with stochastic constraints, Mitra et al. (2018) propose an enhancement model introducing metadata on market sentiment.

In the last few years, a new point of view for the application of stochastic dominance together with enhancement indexation has been introduced. Since the requirement of stochastic dominance in the selected portfolio may be quite restrictive and leading to unfeasible solutions, recent papers (such as Sharma, Agrawal, and Mehra (2017), Bruni et al. (2014)) try to use some relaxed forms of this constraint. Another feature recently investigated by researchers is the attempt of including risk in such a kind of strategies (see Goel, Sharma, and Mehra (2018), Sehgal and Mehra (2019)).

The contribution of this work is the formulation of a linear stochastic dominance enhanced index strategy. Among all the possible dominating portfolios, the aim of this work is to select the one able to mimic the behavior of the benchmark for a given quantile and at the same time with an outperforming minimum level.

The introduction of the quantile asymmetric dispersion measure for the index tracking problem is an important step in the construction of replicating portfolios leading to a linear programming formulation suitable for risk management interpretation (see among the others Wu and Xiao (2002),Meligkotsidou, Vrontos, and Vrontos (2009)).

The enhanced index problem is achieved via first and second order stochastic dominance constraints. Because of these features the proposed model can be applied to large portfolio considering transaction costs and turnover constraints.

The proposed portfolio model is run by constructing rolling strategies and it is tested for several quantile confidence level. The proposed model shows the ability to mimic the benchmark returns with significant extra-performances.

This chapter is organized as follows: In the next section it is introduced the index tracking portfolio problem. Section 3 discusses the proposed enhanced index tracking problem. In Section 4, the empirical analysis are discussed in the static and rolling cases. In the last Section, the obtained results are briefly summarized.

2.2 Index tracking problem: a quantile regression approach

Quantile regression (QR) has been introduced by Koenker and Bassett in 1978 Koenker and Bassett Jr (1978) with the aim to identify a new class of estimators in order to overcome the problems highlighted by using traditional methods (such as least squares estimators). This methodology shows a better ability to describe phenomena characterized by heavy tails without imposing any distributional assumptions. Let $Y \in \mathbb{R}^T$ be the log-return of equity index, the benchmark with realization y_i (for $i = 1, \ldots, T$) , $R = \{r_1, r_2, \ldots, r_N\}$ be the random vector of its N components with $r_i \in \mathbb{R}^T$. Thus, $X = R\beta$ is the portfolio's return and $\beta \in \mathbb{R}^N$ is the vector of portfolio weights. The tracking error (TE) is defined as the vector $\varepsilon = R\beta - Y$, with $\varepsilon \in \mathbb{R}^T$. There are several ways to build an index tracking portfolio since portfolio managers have different constraints and restrictions. A general benchmark tracking problem can be formulated as follows:

$$\begin{aligned} \min_{\beta} \quad & \sigma(X - Y) \\ \text{s.t.} \quad & \sum_{n=1}^{N} \beta_n = 1 \\ & \mathbb{E}[X] - \mathbb{E}[Y] \geq K^* \\ & lb \leq \beta_n \leq ub \qquad \forall n = 1, \ldots, N \end{aligned} \tag{18}$$

where σ is a dispersion measure generated from a given probability metric Stoyanov, Rachev, and Fabozzi (2008). The first constraints impose to invest all the available wealth. The second is related to institutional policy and defines the minimum guaranteed return level K^*. Finally, the last constraint bounds the upper (ub) and lower (lb) value of the portfolio weights. Three dispersion measures are broadly used: the mean absolute deviation (TEMAD), the downside mean semideviation (TEDMS) and the tracking error volatility (TEV).

The problem (18), can be re-conducted to a LP problem when it is considered TEMAD or TEDMS as dispersion measures (see Mansini, Ogryczak, and Speranza (2003)) while the index tracking portfolio obtained with the TEV is a quadratic programming problem. The three measures present some drawbacks and theoretical lacks. In particular, the TEMAD is a symmetric measure with an equal weight for positive and negative ε_t while investors have different preferences and they show a diverse risk profile according to their aversion to negative events. The TEDMS is clearly an asymmetric measure and it is suitable to capture only the downside risk but can lead to portfolios with an intrinsic higher risk. The TEV is the most used measure, among the three here presented, it represents the errors' variance and it is forward-looking oriented. One interpretation of this measure is related to VaR Jorion (2003). However, this measure is still symmetric with respect to the error mean and it takes into account the quadratic variation of the difference between portfolio and benchmark returns.

Our aim is to overcome these limitations considering quantile regression to build a dispersion measure of the tracking error called **TEQR** (Tracking Error Quantile regression). It relates how the quantile of the dependent variable varies with the independent variable.

Definition of TEQR

Let $\varepsilon_t = y_t - \sum_{n=1}^{N} r_{t,n}\beta_n$ *be the difference between portfolio and benchmark returns at time* t *and* $\xi \in \mathbf{R}$ *, the* ***tracking error quantile regression*** *(TEQR) at given* τ *is:*

$$\text{TEQR} \quad \sigma(\varepsilon, \xi|\tau) = \tau \sum_{t=1}^{T} |\varepsilon_t - \xi|_+ + (1-\tau) \sum_{t=1}^{T} |\varepsilon_t - \xi|_- \tag{19}$$

where $|\varepsilon_t|_+ = max(0, \varepsilon_t)$ and $|\varepsilon_t|_- = max(0, -\varepsilon_t)$.

In (19), the first term is the sum of the positive residuals while the second term is the sum of negative residuals. ξ is a variable that could be seen, in the minimization of the TEQR and according with Rockafellar and Uryasev (2013), as the VaR of residuals. $|\varepsilon_t|_+$ are the cumulative errors related to the observations that lie above the regression line and they receive a weight of τ, while $|\varepsilon_t|_-$ are the cumulative ones of the observations that lie below the regression line and they receive a weight of $(1-\tau)$.

From these definition it is evident that the asymmetry property of the tracking error quantile regression (19) is linked to the value of the selected τ. It gives a different weight to positive and negative tracking errors and it also represents an aversion risk coefficient. As underlined by Koenker and Bassett Jr (1978), the quantile regression problem does not present a close form solution and its solution is the result of a minimization problem.

TEQR is a relative deviation metric since it satisfies the following properties presented in Stoyanov et al. (2008).

- $\tilde{P1}$. $\nu(X,Y) \geq 0$ and $\nu(X,Y) = 0$ if and only if $X \overset{\text{a.s.}}{=} Y$;
- $P3$. $\nu(X,Y) \leq \nu(X,Z) + \nu(Z,Y) \quad \forall \quad X, Y, Z$;
- $\tilde{P4}$. $\nu(X+Z, Y+Z) = \nu(X,Y) \quad \forall \quad X, Y, Z$;
- $P5$. $\nu(X+c_1, Y+c_2) = \nu(X,Y) \quad \forall \quad X, Y$ and constants c_1, c_2;
- $P6$. $\nu(aX, aY) = a^s\nu(X,Y) \quad \forall \quad X, Y, a, s \geq 0$

Fixing the value of τ and given these properties, through the Proposition 1 and 2 in Stoyanov et al. (2006) it can be argued that the tracking error quantile regression can be seen as a translation invariant metric and it is also a deviation measure in the sense of the definition presented by Rockafellar, Uryasev, and Zabarankin (2006).

The tracking error problem with the TEQR measure can be formulated as a linear problem using positive auxiliary variables $u_t and \nu_t$:

$$
\begin{aligned}
\min_{\beta,u,\nu} \quad & \sum_{t=1}^{T} \tau u_t + (1-\tau)\nu_t \\
\text{s.t.} \quad & r_t\beta - u_t + \nu_t - \xi = y_t \quad \forall t = 1,\dots,T \\
& \sum_{n=1}^{N} \beta_n = 1 \\
& \mathbb{E}[X] - \mathbb{E}[Y] \geq K^* \\
& u_t,\, \nu_t \geq 0 \qquad\qquad \forall t = 1,\dots,T \\
& lb \leq \beta_n \leq ub \qquad\quad \forall n = 1,\dots,N
\end{aligned}
\tag{20}
$$

2.3 Stochastic dominance constraints

The relation of stochastic dominance is one of the fundamental concepts of the decision theory Levy (1992).

The first degree relation carries over to expectations of monotone utility functions, and the second degree relation to expectations of concave non-decreasing utility functions. The first order definition of stochastic dominance (FSD) gives a partial order in the space of real random variables Levy (1992); Bawa (1978).

Let X and Y be RVs of the returns of two financial portfolios. Then, in the stochastic dominance approach, they are compared through some performance functions constructed from their distributions. For a real random variable X, its first performance function is defined as the right-continuous cumulative distribution function of X:

$$F_X(\xi) = \mathbb{P}(X \leq \xi) \quad for\ \xi \in \mathbb{R} \tag{21}$$

A random return X is said to stochastically dominate another random return Y in the first order sense, denoted $X \underset{(1)}{\geq} Y$, if

$$F_X(\xi) \leq F_Y(\xi) \quad for\ \xi \in \mathbb{R} \tag{22}$$

More important from the portfolio point of view is the notion of second-order dominance (SSD). It is one of the most debated topic in financial portfolio selection, due to its connection to the theory of risk-averse investor behavior and tail risk minimization Bawa (1975).

It is equivalent to this statement: a random variable X dominates the random variable Y if $\mathbb{E}[u(X)] \geq \mathbb{E}[u(Y)]$ for all non-decreasing concave functions $u(\cdot)$ for which these expected values are finite. Thus, no risk-averse decision maker will prefer a portfolio with return rate Y over a portfolio with return rate X Lozza, Shalit, and Fabozzi (2013). The second performance function $F^{(2)}$ is given by area below the distribution function F:

$$F_X^{(2)} = \int_{-\infty}^{z} F_X(z)dz \qquad \forall \qquad z \in \mathbb{R} \tag{23}$$

and defines the weak relation of the second-order stochastic dominance. Which is, random return X stochastically dominates Y in the second order, denoted $X \underset{(2)}{\geq} Y$, if

$$F_X^{(2)}(z) \leq F_Y^{(2)}(z) \qquad \forall \qquad z \in \mathbb{R} \tag{24}$$

Changing the order of integration, the ordering $X \underset{(2)}{\geq} Y$ is equivalent to the expected shortfall Lozza, Shalit, and Fabozzi (2013); Ogryczak and Ruszczyński (1999):

$$F_X^{(2)}(z) = \mathbb{E}[(z - X)_+] \qquad \forall \qquad z \in \mathbb{R} \tag{25}$$

In this case, the function $F_X^{(2)}(\xi)$ is continuous, convex, nonnegative and non-decreasing. It is well defined for all random variables X with finite expected value. The introduction of stochastic dominance in the index tracking portfolio problem is also presented in Roman, Mitra, and Zverovich (2013) where they apply a second order stochastic dominance strategy in a multi-objective linear problem minimizing the tail risk of the strategy. Bruni et al. (2014) propose a linear bi-objective optimization approach to maximize the average excess return minimizing the risk. Then, they investigate the theoretical condition to guarantee the existence of an enhanced index portfolio.

Computational tractable and technological solvable portfolio optimization models which apply the concept of FSD or SSD were proposed by Lozza, Shalit, and Fabozzi (2013); Kuosmanen (2004).

In this formulation it is followed the methodology to linearize FSD and SSD presented in Kopa (2010) and Kuosmanen (2004). They reached the goal of linearizing the objective function through the introduction of a permutation matrix. Let $P = \{p_{r,c}\}$ a permutation matrix with $p_{r,c} = 0, 1$ such that $\sum_{r=1}^{T} p_{r,c} = 1$ for $c = 1, \ldots, T$ and $\sum_{c=1}^{T} p_{r,c} = 1$ for $r = 1, \ldots, T$. Then portfolio $X = R\beta$, with $\beta \in \mathbb{R}^N$ and R is the $T \times N$ matrix of equiprobable asset's returns, dominates portfolio Y in a first order sense if and only if:

$$\begin{aligned}
& X \geq PY \\
& \sum_{r=1}^{T} p_{r,c} = 1 \quad \forall c = 1,\dots,T \\
& \sum_{c=1}^{T} p_{r,c} = 1 \quad \forall r = 1,\dots,T \\
& p_{r,c} \in 0,1 \qquad \forall r = 1,\dots,T;\ \ \forall c = 1,\dots,T
\end{aligned} \tag{26}$$

Kopa (2010) and Kuosmanen (2004) also propose a linear formulation of the second order stochastic dominance. Assuming that the returns have a discrete joint distribution with realizations $x_t,\ t = 1,\dots,T$ having the same probability, then $X \underset{(2)}{\geq} Y$ in the second order stochastic dominance sense if and only if it exists a double stochastic matrix $Z = \{z_{r,c}\}$ with $z_{r,c} \in [0,1]$ such that

$$\begin{aligned}
& X \geq ZY \\
& \sum_{r=1}^{T} z_{r,c} = 1 \quad \forall c = 1,\dots,T \\
& \sum_{c=1}^{T} z_{r,c} = 1 \quad \forall r = 1,\dots,T \\
& 0 \leq z_{r,c} \leq 1 \quad \forall r = 1,\dots,T;\ \ \forall c = 1,\dots,T
\end{aligned} \tag{27}$$

2.4 Enhanced Indexing Problem with Stochastic Dominance Constraints

A realistic formulation to solve the enhanced index benchmark tracking problem should consider the introduction of some relevant features to real world investment strategies. This new formulation takes into account a linear penalty objective function to reduce the portfolio turnover and risk management duties. Moreover, although the introduction of stochastic dominance constraints enhances the benchmark tracking model, its formulation strongly increases the dimensionality and the computational complexity of the problem. In particular, considering the methodologies proposed by Kopa (2010) and Kuosmanen (2004) it is possible to solve the problem efficiently.

The enhanced index benchmark tracking problem is solved considering the minimization of a dispersion measure of the tracking error, the TEQR (19), which could be formulated as linear program. To enhance the performance in the risk minimization, first and second order stochastic dominance constraints are introduced following the formulations (26) and (27).

Additionally, transaction costs are introduced in the objective function to improve real-life performances of the portfolio. Let $tc = tc^{+} + tc^{-}$be the transaction total cost where: $tc^{-} = \alpha \cdot \omega^{-}$ and $tc^{+} = \alpha \cdot \omega^{+}$ are the fees for trading assets. With α is the proportional constant for transaction costs, ω^{+} and ω^{-} are the changes in portfolio weights and $\omega = \sum(\omega^{+} + \omega^{-})$. The enhanced indexation benchmark tracking problem with FSD constraints is defined as:

$$\begin{aligned}
\min_{\beta,u,\nu,p,\omega^{+},\omega^{-}} \quad & \sum_{t=1}^{T} \tau u_t + (1-\tau)\nu_t + tc\Delta\omega \\
\text{s.t.} \quad & r_t\beta - u_t + \nu_t = y_t && \forall t = 1,\dots,T \\
& \sum_{n=1}^{N} \beta_n = 1 \\
& \mathbb{E}[X] - \mathbb{E}[Y] \geq K^{*} \\
& \omega_n^{+} - \omega_n^{-} = \beta_n - \beta_n^{old} && \forall n = 1,\dots,N \\
& \textstyle\sum_n |\beta_n - \beta_n^{old}| \leq \theta && n = 1,\dots,N \\
& X \geq PY \\
& \sum_{r=1}^{T} p_{r,c} = 1 && \forall c = 1,\dots,T \\
& \sum_{c=1}^{T} p_{r,c} = 1 && \forall r = 1,\dots,T \\
& p_{r,c} \in \{0,1\} && \forall r,c = 1,\dots,T \\
& u_t,\, \nu_t \geq 0 && \forall t = 1,\dots,T
\end{aligned} \tag{28}$$

The solution to this problem is the portfolio optimal for all insatiable investors. As discussed in Jarrow (1986), the existence of a portfolio that stochastically dominates the index in a first order sense is equivalent to the presence of arbitrage.
The enhanced index benchmark tracking problem (28) is a mixed-integer linear programming since the permutation matrix P is composed by binary variables. It can be noticed how the dimensionality of this problem quadratically increases together with the number of observation T.

The other proposed model is based on second order stochastic dominance with the introduction a double stochastic matrix Z. Thus, the enhanced indexation benchmark tracking problem with

SSD constraints can be formulated as:

$$
\begin{aligned}
\min_{\beta,u,\nu,p,\omega^+,\omega^-} \quad & \sum_{t=1}^{T} \tau u_t + (1-\tau)\nu_t + tc\Delta\omega & \\
\text{s.t.} \quad & r_t\beta + u_t - \nu_t = y_t & \forall t = 1,\dots,T \\
& \sum_{n=1}^{N} \beta_n = 1 & \\
& \mathbb{E}\left[X\right] - \mathbb{E}\left[Y\right] \geq K^* & \\
& \omega_n^+ - \omega_n^- = \beta_n - \beta_n^{old} & \forall n = 1,\dots,N \\
& \sum\nolimits_n |\beta_n - \beta_n^{old}| \leq \theta & n = 1,\dots,N \\
& X \geq ZY & \\
& \sum_{r=1}^{T} z_{r,c} = 1 & \forall c = 1,\dots,T \\
& \sum_{c=1}^{T} z_{r,c} = 1 & \forall r = 1,\dots,T \\
& 0 \leq z_{r,c} \leq 1 & \forall r,c = 1,\dots,T \\
& u_t,\ \nu_t \geq 0 & \forall t = 1,\dots,T
\end{aligned}
\tag{29}
$$

Differently from the previous enhanced index problem with first order stochastic dominance constraints, this formulation is a linear programming and could be efficiently solved also when the computational complexity increases together with the number of observations.

2.5 Empirical Application

In this section the empirical results of the work are analyzed. The imposition of stochastic dominance's constraint tends to be very restrictive for the feasibility area (in first order and with monthly re-balancing there are very few feasible solutions), moreover computational complexity of the problem is greater for first order models (caused by the presence of binary variable in the permutation matrix). Considering these computational issues, and the fact that the most interesting constraint in finance is the Second order Stochastic Dominance (SSD) one, the FSD model is not investigated here. In particular, the model is tested in the in-sample and out-of-sample context in order to catch the real application of our model. The data set collected in order to perform the analysis is made by the adjusted closing prices of a selection of constituents the Standard and Poor 500 Index through which an artificial benchmark portfolio and an investment strategy are built coherently with our model. It has chosen the constituents that remain in the index for the entire analysis period from 10th March 2010 to 22nd May 2019.

The calibration has been done over the first set of historical daily observations. The portfolio rebalancing is assumed every 21 days (i.e. $T = 21$, one month) and the calibration window moves over the same period. So, every empirical computation is made with 100 different optimization cycles. For this analysis, the benchmark is an equally weighted portfolio of the components of the *Standard* & *Poor* 500 and the investment portfolio is built over an assets' pre-selection of the first 50 assets with lowest volatility . The strategy is tested at three different quantiles, $\tau = [0.05, 0.5, 0.95]$. It is imposed an enhancement parameter with $K^* = 0.0001$, a transaction cost factor α equal to 2% of the changing wealth invested in each asset. Computations are made using GAMS and MATLAB with GUROBI as solver.

2.5.1 Portfolio wealth

As it is possible to observe from Figure 10 and Figure 11, the wealth generated by the Enhanced Indexing strategy with Stochastic Dominance Constraints is always above the one related to the benchmark for each τ tested. A nice feature could be found evaluating the ranking among strategies: in the in-sample analysis, the choice of an higher τ is preferable. In the out-of-sample analysis, the choice is flipped: the intermediate τ performs better. It has been found just one infeasible solution: in this occurrence it has been preserved the previous portfolio composition.

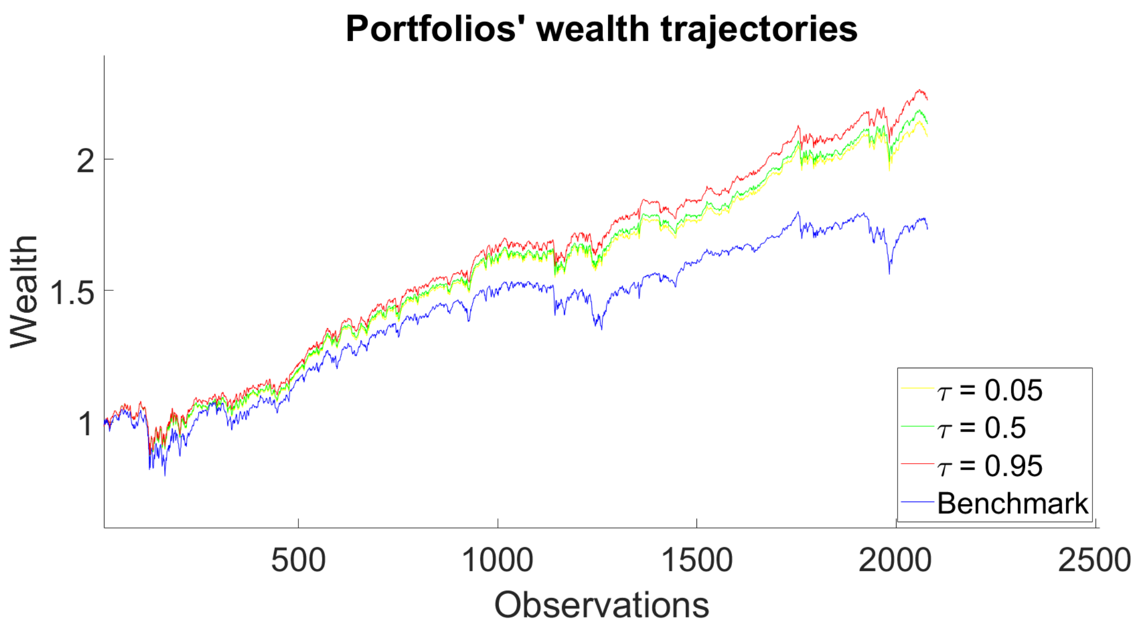

Figure 10: Evolution of wealth generated by the portfolios and benchmark over an horizon of 10 years with in-sample data.

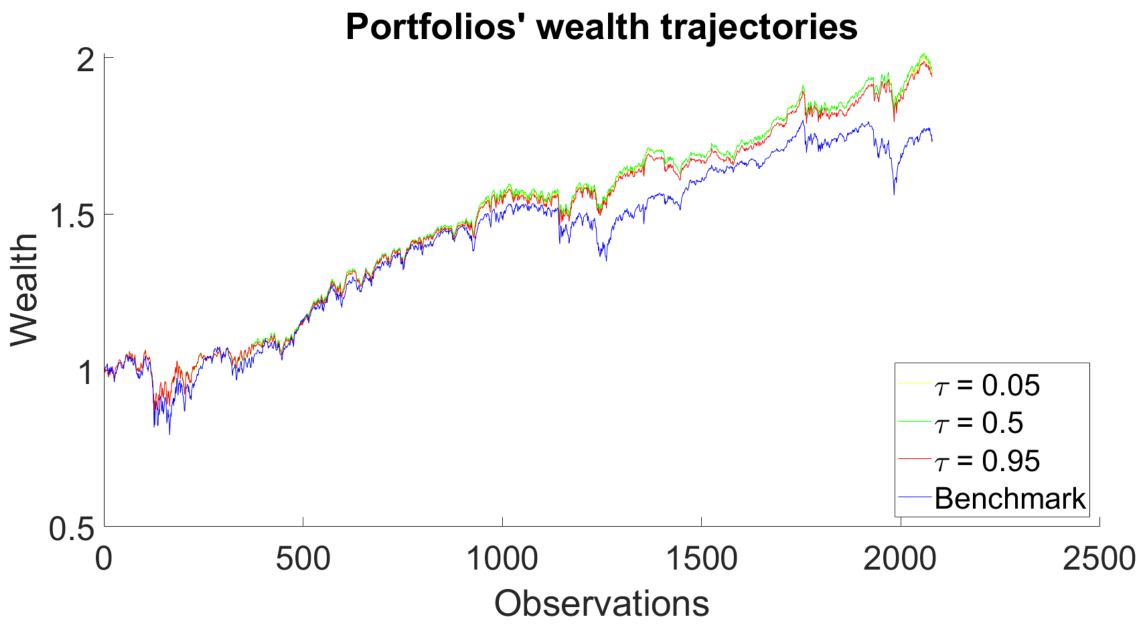

Figure 11: Evolution of wealth generated by the portfolios and benchmark over an horizon of 10 years with out-of-sample data.

2.5.2 Accumulated cumulative empirical functions

In this section it is presented, with a graphical interpretation, the fulfillment of stochastic dominance constraints. Even though in the out-of-sample analysis the distance between the two functions (Benchmark and portfolio $\tau = 0.95$) is smaller than the one in the in-sample case, it is possible to see that the investment line is always above the benchmark one and this implies SSD. It is useful to point out that the graphical interpretation is flipped with respect to the most common one, this is because it is done with empirical distributions and in this case supports are different.

2.5.3 Portfolio concentration

One of the main problems affecting tracking error models is the cardinality of portfolios. It is common that the optimization leads to extremely fragmented portfolios, composed by a large number of assets and therefore with a difficult real-life application. In literature several authors observed the impact of stochastic dominance constraints on cardinality, they underlined that such kind of strategies naturally tend to reduce the number of active securities, or in any case in increasing the degree of concentration of the portfolio. In this case it is decided not to impose an explicit cardinality constraint, it has just been taken in account the effect of the SSD constraint. On the other side, a turnover constraint is imposed in order to keep the investment more stable and adherent to the real world. As it is possible to see in Figure 14 and Figure 15, portfolios are composed by many assets with low weight and less securities with high weight. An interesting

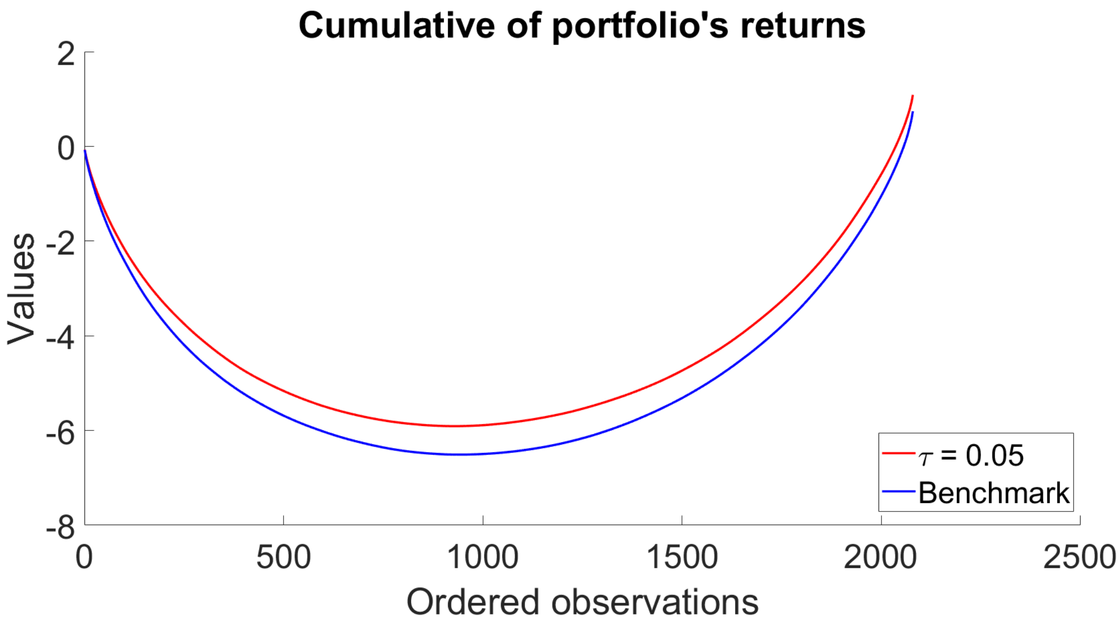

Figure 12: Accumulated returns of the benchmark and of the portfolio with $\tau = 0.95$ in the in-sample analysis

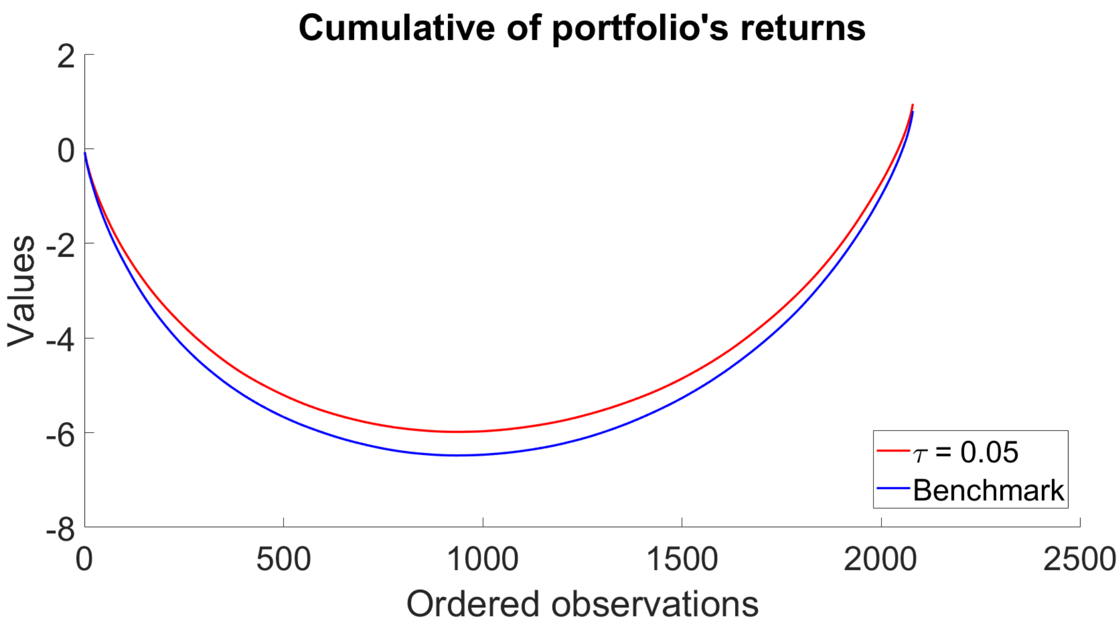

Figure 13: Accumulated returns of the benchmark and of the portfolio with $\tau = 0.95$ in the out-of-sample analysis

feature is that assets with higher weights are the same along the entire investment horizon, but with different proportions. This feature is also observable in the comparison among different τ.

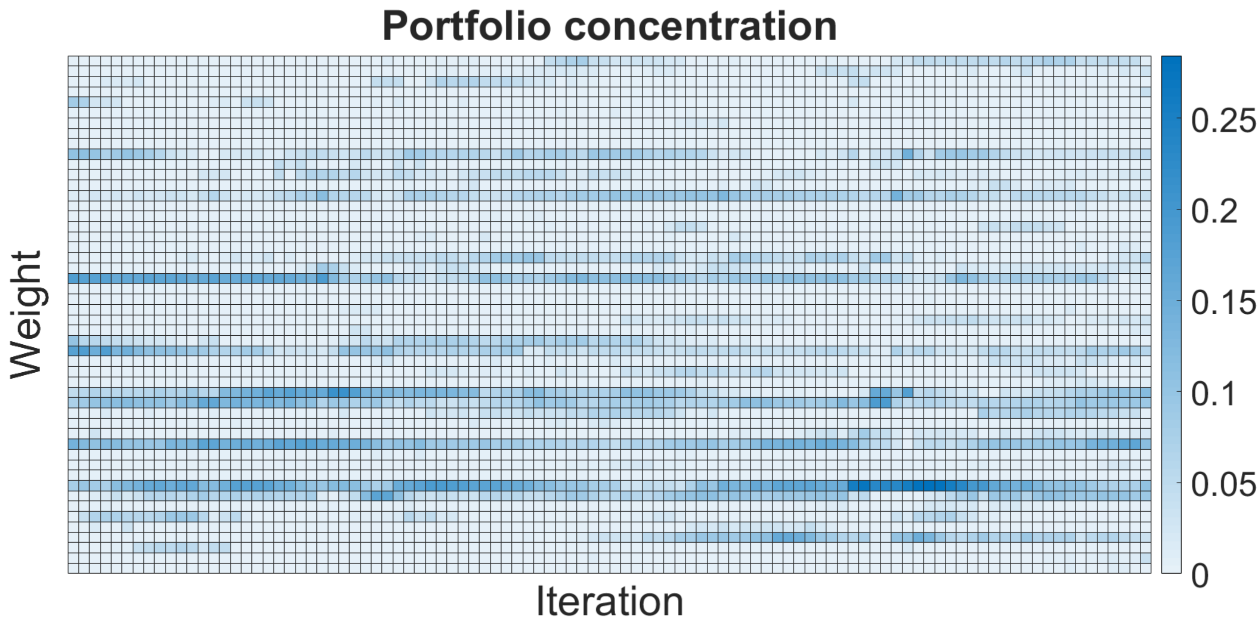

Figure 14: Portfolio concentration with respect to each calibration window.

2.5.4 Statistical analysis

In the following tables it is possible to observe a numerical evidence of the intuition arose by looking at the trajectories. In the in-sample analysis, the best choice, both in terms of performance and in terms of volatility, is the one with higher τ. This choice lead the model to heavy penalization of the "negative" errors. In the out-of-sample analysis, the best choice is different, in terms of volatility the choice is the same, but when looking at the performance it is preferable to chose lower τ.

Numerically, the in-sample analysis return a portfolio with a final wealth between 21,6% and 112,8% higher with respect to the benchmark, with a standard deviation around 0.88% compared to the 0.99% of the equally weighted portfolio. In the out-of-sample analysis it is reached a portfolio with final wealth between 15,6% and 26,2% higher with respect to the benchmark, and a standard deviation similar to the in-sample one.

2.6 Conclusions

This chapter presents a semi-passive investment strategy. A new dispersion measure, TEQR, is proposed for asset management. The resulting portfolio maintains the evolution of the tracked benchmark with the addition of an enhancement parameter K^*. The investment portfolio

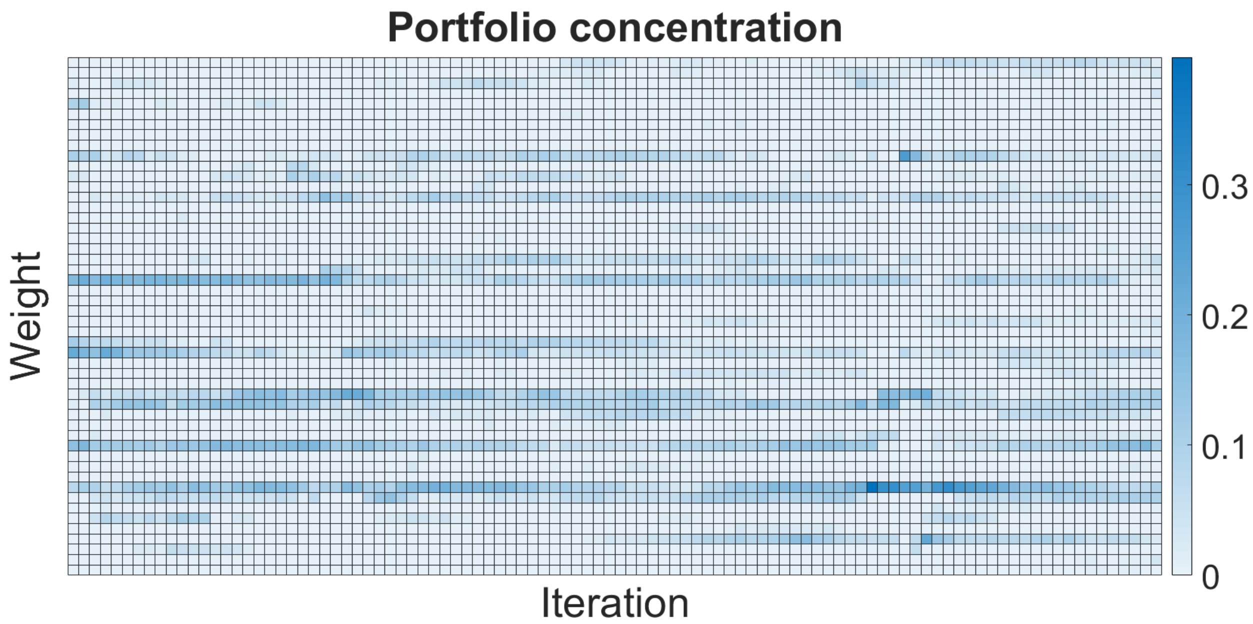

Figure 15: Portfolio concentration with respect to each calibration window.

	Mean	**Std**	**Kurtosis**	**Skewness**
$\tau = 0.05$	0.0004882	0.0087945	8.0174375	-0.4493619
$\tau = 0.5$	0.0005362	0.0087645	8.1053605	-0.4592499
$\tau = 0.95$	0.0006086	0.0087359	8.1917696	-0.4797739

Table 3: Statistics computed over in-sample data

	Mean	**Std**	**Kurtosis**	**Skewness**
$\tau = 0.05$	0.0004727	0.0087975	8.0047749	-0.4703994
$\tau = 0.5$	0.0004771	0.0087800	8.0946631	-0.4850888
$\tau = 0.95$	0.0004581	0.0087565	8.1128079	-0.5046264

Table 4: Statistics computed over out-of-sample data

is built with the requirement of second order stochastic constrain. Both in in-sample and in out-of-sample analysis, the proposed portfolio performs better than the benchmark. In the outputs of the model there is the optimal ξ which can be interpreted as the VaR of the errors between benchmark and the selected portfolio. This chapter shows good features of the proposed strategy and this could be a starting point for future researches. Possible developments of this work could be represented by the research of pre-selection strategy for the assets available for investing and in the analysis of ξ as determining factor for the choice of the optimal portfolio.

2.7 Annex: Choosing the calibrating window

In order to choose the length of the calibrating window for this model it has been performed a sensitivity analysis over a smaller problem. It is tested, in the out-of-sample the reactivity of the model over the first 50 calibrating windows. It is studied the different evolution of the errors' VaR ξ (as it is possible to see in figure 16, where darker lines are related to wider calibrating windows. At the top it is possible to find ξ with $\tau = 0.95$, at the bottom with $\tau = 0.05$ and in the middle with $\tau = 0.5$. The plot presents just four line (6M, 7M, 1Y and 3Y) in order to make the graph more readable. The blue line represents the dynamic of the benchmark portfolio) in order to investigate the utility of this parameter in portfolio selection. It is computed absolute mean and standard deviation for each quantile and calibrating window [6 months, 7 months, 12 months and 15 months]. The model is not tested for window smaller than six months because the feasibility set of the problem became even smaller ad it is very hard to find acceptable solutions. Firstly it can be seen that with larger windows, the model provides more stable trajectories of the ξ at every level of τ.
Looking at the variance and absolute mean of our portfolios as in Table 5, it is possible to argue that the standard deviation remains stable, but the wider is the calibrating window (until two years), the better are the results with respect to the mean. Since the enhancement between one year and two years is more and more reduced, but on the other hand the computational time increase largely. So, the optimal calibrating window is set at 315 days.

	Standard Deviation			**Absolute mean**		
	$\tau = 0.05$	$\tau = 0.5$	$\tau = 0.95$	$\tau = 0.05$	$\tau = 0.5$	$\tau = 0.95$
6M	0.00780	0.00781	0.00792	0.00047	0.00045	0.00058
7M	0.00779	0.00781	0.00791	0.00049	0.00047	0.00057
12M	0.00770	0.00778	0.00779	0.00056	0.00057	0.00061
15M	0.00782	0.00783	0.00795	0.00060	0.00061	0.00062
36M	0.00794	0.00792	0.00803	0.00060	0.00059	0.00059

Table 5: Statistics over strategies computed to calibrate the lenght of the window

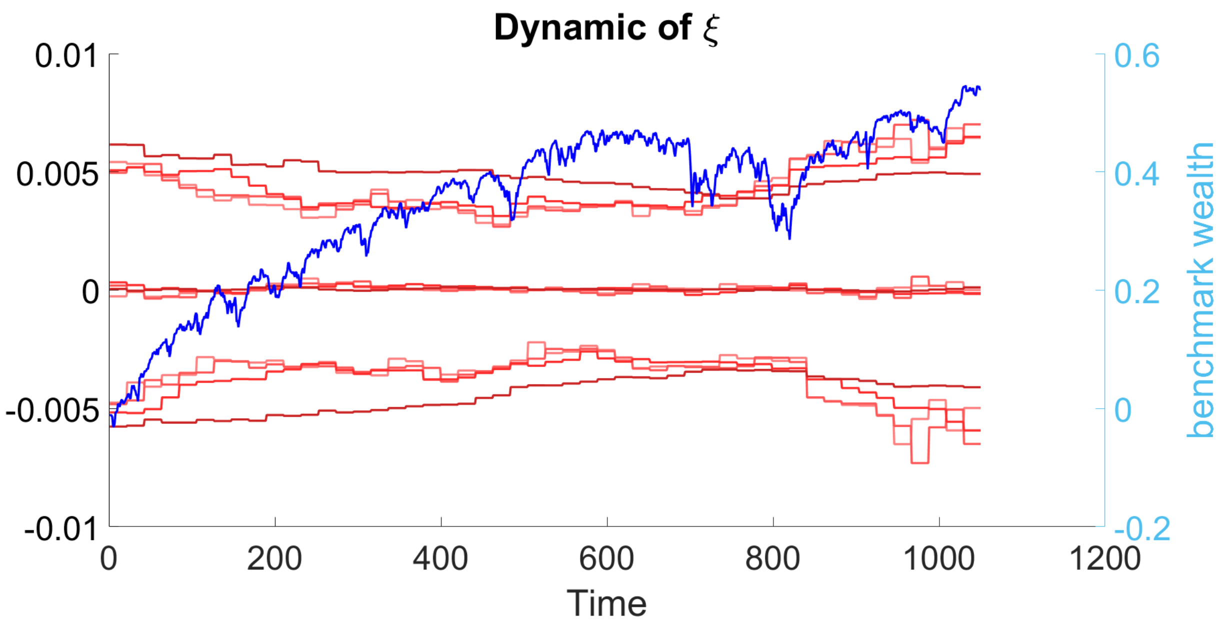

Figure 16: Dynamics of different ξ

.

List of Figures

1 Interpretation of the relevance given to errors in quantile regression with $\tau = 0.95$. 10
2 Interpretation of the quantile regression line with $\tau = 0.95, \tau = 0.5, \tau = 0.05$. . . 11
3 Graphical interpretation of the Markov tree 18
4 Dynamic Sharpe Ratio (DSR) and Stochastic Benchmark ratio (SBR) computed on a portfolio of three assets 24
5 Comparison among cumulated wealth with Sharpe type performance strategies . 26
6 Comparison among conditional dynamic Sharpe type strategies 27
7 Comparison of conditional stochastic benchmark strategies 27
8 Joint distributions (wealth and variance at time T) of HL and HH optimal DSR portfolios 30
9 Joint distributions (wealth and variance at time T) of LL and LH optimal DSR portfolios 30
10 Evolution of wealth generated by the portfolios and benchmark over an horizon of 10 years with in-sample data. 42
11 Evolution of wealth generated by the portfolios and benchmark over an horizon of 10 years with out-of-sample data. 43
12 Accumulated returns of the benchmark and of the portfolio with $\tau = 0.95$ in the in-sample analysis 44
13 Accumulated returns of the benchmark and of the portfolio with $\tau = 0.95$ in the out-of-sample analysis 44
14 Portfolio concentration with respect to each calibration window. 45
15 Portfolio concentration with respect to each calibration window. 46
16 Dynamics of different ξ 49

List of Tables

1 Stochastic dominance among strategies and Benchmark. 28
2 Average ex-ante statistics of the forecasted final wealth and volatility for the optimal portfolio of the conditional Sharpe type strategies 29
3 Statistics computed over in-sample data 46
4 Statistics computed over out-of-sample data 46
5 Statistics over strategies computed to calibrate the lenght of the window 48

References

Anderson, Gordon (1996). "Nonparametric tests of stochastic dominance in income distributions". In: *Econometrica: Journal of the Econometric Society*, pp. 1183–1193.

Angelelli, E. and S. Ortobelli Lozza (2009). "American and European portfolio selection strategies: The Markovian approach". In: *Financial hedging* 5, pp. 119–152.

Artzner, P. et al. (1999). "Coherent measures of risk". In: *Mathematical finance* 9(3), pp. 203–228.

Barrett, Garry F and Stephen G Donald (2003). "Consistent tests for stochastic dominance". In: *Econometrica* 71(1), pp. 71–104.

Bawa, V. S. (1975). "Optimal rules for ordering uncertain prospects". In: *Journal of Financial Economics* 2(1), pp. 95–121.

Bawa, V. S. (1978). "Safety-first, stochastic dominance, and optimal portfolio choice". In: *Journal of Financial and Quantitative Analysis* 13(2), pp. 255–271.

Bean, N. G., N. Kontoleon, and P. G. Taylor (2008). "Markovian trees: properties and algorithms". In: *Annals of Operations Research* 160(1), pp. 31–50.

Bollerslev, T. (1986). "Generalized autoregressive conditional heteroskedasticity". In: *Journal of econometrics* 31(3), pp. 307–327.

Bruni, R. et al. (2014). "A linear risk-return model for enhanced indexation in portfolio optimization". In: *OR Spectrum*, pp. 1–25.

Canakgoz, N. A. and J. E. Beasley (2009). "Mixed-integer programming approaches for index tracking and enhanced indexation". In: *European Journal of Operational Research* 196(1), pp. 384–399.

Çanakoğlu, Ethem and Süleyman Özekici (2009). "Portfolio selection in stochastic markets with exponential utility functions". In: *Annals of Operations Research* 166(1), p. 281.

Cogneau, Philippe and Georges Hübner (2009a). "The (more than) 100 Ways to Measure Portfolio Performance - Part 1: standardized risk-adjusted measures. Journal of Performance Measurement, 13(Summer), 56-71". In: *Journal of Performance Measurement* 13, pp. 56–71.

Cogneau, Philippe and Georges Hübner (2009b). "The (more than) 100 Ways to Measure Portfolio Performance - Part 2: Special Measures and Comparison". In: *Journal of Performance Measurement* 14(1), pp. 56–69.

D'Amico, G. and G. Di Biase (2009). "Dynamic Concentration /Inequality Indices of Economic Systems". In: *In proceedings of the International Conference "Recent Advances in Applied Mathematics" (C.A. Bolucea, V. Mladenov, E. Pop, M. Leba, N. Mastorakis, Eds.)* Pp. 312–316.

D'Amico, G et al. (2009). "Semi-Markov backward credit risk migration models compared with Markov models". In: *3RD International Conference on Applied Mathematics, Simulation, Modelling*, pp. 112–115.

D'Amico, G et al. (2010). "Semi-Markov Backward Credit Risk Migration. Models: a Case Study". In: *International Journal of Mathematical Models and Methods in Applied Sciences* 4(1), pp. 82–92.

Davidson, Russell and Jean-Yves Duclos (2000). "Statistical inference for stochastic dominance and for the measurement of poverty and inequality". In: *Econometrica* 68(6), pp. 1435–1464.

Dentcheva, D. and A. Ruszczyński (2006). "Portfolio optimization with stochastic dominance constraints". In: *Journal of Banking & Finance* 30(2), pp. 433–451.

Duan, J. C. and J. G. Simonato (2001). "American option pricing under GARCH by a Markov chain approximation". In: *Journal of Economic Dynamics and Control* 25(11), pp. 1689–1718.

Elliott, R. J. and T. K. Siu (2010). "On risk minimizing portfolios under a Markovian regime-switching Black-Scholes economy". In: *Annals of Operations Research* 176(1), pp. 271–291.

Engle, R. F. (1982). "Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation". In: *Econometrica: Journal of the Econometric Society*, pp. 987–1007.

Fábián, C. I., G. Mitra, and D. Roman (2011). "Processing second-order stochastic dominance models using cutting-plane representations". In: *Mathematical Programming* 130(1), pp. 33–57.

Fábián, C. I. et al. (2011). "An enhanced model for portfolio choice with SSD criteria: a constructive approach". In: *Quantitative Finance* 11(10), pp. 1525–1534.

Fu, Y. H. et al. (2015). "Portfolio optimization with transaction costs: a two-period mean-variance model". In: *Annals of Operations Research* 233(1), pp. 135–156.

Glosten, L. R., R. Jagannathan, and D. E. Runkle (1993). "On the relation between the expected value and the volatility of the nominal excess return on stocks". In: *The journal of finance* 48(5), pp. 1779–1801.

Goel, A., A. Sharma, and A. Mehra (2018). "Index tracking and enhanced indexing using mixed conditional value-at-risk". In: *Journal of Computational and Applied Mathematics* 335, pp. 361–380.

Guastaroba, G. et al. (2016). "Linear programming models based on Omega ratio for the enhanced index tracking problem". In: *European Journal of Operational Research* 251(3), pp. 938–956.

Hadar, Josef and William R Russell (1969). "Rules for ordering uncertain prospects". In: *The American economic review* 59(1), pp. 25–34.

Hanoch, Giora and Haim Levy (1969). "The efficiency analysis of choices involving risk". In: *The Review of Economic Studies* 36(3), pp. 335–346.

Ibe, Oliver (2014). *Fundamentals of applied probability and random processes.* Academic Press.

Jarrow, R. (1986). "The relationship between arbitrage and first order stochastic dominance". In: *The Journal of Finance* 41(4), pp. 915–921.

Jorion, P. (2003). "Portfolio optimization with constraints on tracking error". In: *Financial Analysts Journal* 59(5), pp. 70–82.

Karamata, Jovan (1932). "Sur une inégalité relative aux fonctions convexes". In: *Publications de l'Institut Mathématique* 1(1), pp. 145–147.

Keating, C. and W. F. Shadwick (2002). "A universal performance measure". In: *Journal of performance measurement* 6(3), pp. 59–84.

Koenker, R. and G. Bassett Jr (1978). "Regression quantiles". In: *Econometrica: Journal of the Econometric Society*, pp. 33–50.

Koenker, Roger and Zhijie Xiao (2002). "Inference on the quantile regression process". In: *Econometrica* 70(4), pp. 1583–1612.

Kopa, M. (2010). "Measuring of second–order stochastic dominance portfolio efficiency". In: *Kybernetika* 46, pp. 488–500.

Kuosmanen, T. (2004). “Efficient diversification according to stochastic dominance criteria”. In: *Management Science* 50(10), pp. 1390–1406.

Leshno, Moshe and Haim Levy (2002). “Preferred by “all” and preferred by “most” decision makers: Almost stochastic dominance”. In: *Management Science* 48(8), pp. 1074–1085.

Levy, H. (1992). “Stochastic dominance and expected utility: survey and analysis”. In: *Management science* 38(4), pp. 555–593.

Lo, A. W. (2004). “The adaptive markets hypothesis”. In: *The Journal of Portfolio Management* 30(5), pp. 15–29.

Lozza, S. O., H. Shalit, and F. J. Fabozzi (2013). “Portfolio selection problems consistent with given preference orderings”. In: *International Journal of Theoretical and Applied Finance* 13(5).

Malkiel, B. G. and E. F. Fama (1970). “Efficient capital markets: A review of theory and empirical work”. In: *The journal of Finance* 25(2), pp. 383–417.

Mansini, R., W. Ogryczak, and M. G. Speranza (2003). “LP solvable models for portfolio optimization: A classification and computational comparison”. In: *IMA Journal of Management Mathematics* 14(3), pp. 187–220.

Markowitz, Harry (1952). “Portfolio selection”. In: *The journal of finance* 7(1), pp. 77–91.

Meade, N and John E Beasley (2011). “Detection of momentum effects using an index outperformance strategy”. In: *Quantitative Finance* 11(2), pp. 313–326.

Meligkotsidou, L., I. D. Vrontos, and S. D. Vrontos (2009). “Quantile regression analysis of hedge fund strategies”. In: *Journal of Empirical Finance* 16(2), pp. 264–279.

Mitra, G. et al. (2018). “Using Market Sentiment to Enhance Second-Order Stochastic Dominance Trading Models”. In: *High-Performance Computing in Finance*, pp. 25–48.

Nelson, Daniel B (1991). “Conditional heteroskedasticity in asset returns: A new approach”. In: *Econometrica: Journal of the Econometric Society*, pp. 347–370.

Ogryczak, W. and A. Ruszczyński (1999). “From stochastic dominance to mean-risk models: Semideviations as risk measures”. In: *European Journal of Operational Research* 116(1), pp. 33–50.

Ortobelli Lozza, S., E. Angelelli, and A. Bianchi (2011). “Financial Application of bivariate Markov processes”. In: *Mathematichal problems in Engineering.*

Ortobelli Lozza, S. and G. Iaquinta (2008). “Markov chain applications to non parametric option pricing theory”. In: *IJCSNS* 8(6), p. 199.

Pratt, J. W. (1978). “Risk aversion in the small and in the large”. In: *Uncertainty in Economics*, pp. 59–79.

Rachev, S. (1991). *Probability metrics and the stability of stochastic models*. Vol. 269. John Wiley & Son Ltd.

Rachev, S. T. and S. Mittnik (2000). *Stable Paretian models in finance*. Vol. 7. John Wiley & Son Ltd.

Rao, Calyampudi Radhakrishna et al. (1973). *Linear statistical inference and its applications*. Vol. 2. Wiley New York.

Robertson, Tim and FT Wright (1981). “Likelihood ratio tests for and against a stochastic ordering between multinomial populations”. In: *The Annals of Statistics*, pp. 1248–1257.

Rockafellar, R. T. and S. Uryasev (2013). “The fundamental risk quadrangle in risk management, optimization and statistical estimation”. In: *Surveys in Operations Research and Management Science* 18(1-2), pp. 33–53.

Rockafellar, R. T., S. Uryasev, and M. Zabarankin (2006). “Generalized deviations in risk analysis”. In: *Finance and Stochastics* 10(1), pp. 51–74.

Rockafellar, R. T., S. P. Uryasev, and M. Zabarankin (2002). “Deviation measures in risk analysis and optimization”. In: *University of Florida, Department of Industrial & Systems Engineering Working Paper, 7*.

Roman, D., G. Mitra, and V. Zverovich (2013). “Enhanced indexation based on second-order stochastic dominance”. In: *European Journal of Operational Research* 228(1), pp. 273–281.

Rothschild, Michael and Joseph E Stiglitz (1970). “Increasing risk: I. A definition”. In: *Journal of Economic theory* 2(3), pp. 225–243.

Schmid, Friedrich and Mark Trede (1996). “Testing for First-Order Stochastic Dominance: A New Distribution-Free Test”. In: *Journal of the Royal Statistical Society: Series D (The Statistician)* 45(3), pp. 371–380.

Sehgal, R. and A. Mehra (2019). “Enhanced indexing using weighted conditional value at risk”. In: *Annals of Operations Research*, pp. 1–30.

Seneta, E. (1996). “Markov and the Birth of Chain Dependence Theory”. In: *International Statistical Review / Revue Internationale De Statistique* 64(3), pp. 255–263.

Sharma, A., S. Agrawal, and A. Mehra (2017). “Enhanced indexing for risk averse investors using relaxed second order stochastic dominance”. In: *Optimization and Engineering* 18(2), pp. 407–442.

Sharpe, W. F. (1994). “The sharpe ratio”. In: *Journal of portfolio management* 21(1), pp. 49–58.

Sortino, F. A. and L. N. Price (1994). “Performance measurement in a downside risk framework”. In: *The Journal of Investing* 3(3), pp. 59–64.

Stiglitz, J. E. (1970). “Review of some aspects of theory of risk bearing by kj arrow”. In: *Econometrica* 38.

Stoyanov, S. et al. (2006). *Relative deviation metrics with applications in finance*. Tech. rep. Department of Probability and Applied Statistics - University of California, Santa Barbara (USA).

Stoyanov, S. V., S. T. Rachev, and F. J. Fabozzi (2008). “Probability metrics with applications in finance”. In: *Journal of Statistical Theory and Practice* 2(2), pp. 253–277.

Stoyanov, S. V. et al. (2008). “Relative deviation metrics and the problem of strategy replication”. In: *Journal of Banking & Finance* 32(2), pp. 199–206.

Tauchen, G. (1986). “Finite state markov-chain approximations to univariate and vector autoregressions”. In: *Economics letters* 20(2), pp. 177–181.

Valladão, D., T. Silva, and M. Poggi (2018). “Time-consistent risk-constrained dynamic portfolio optimization with transactional costs and time-dependent returns”. In: *Annals of Operations Research*, pp. 1–27.

Valle, C. A., D. Roman, and G. Mitra (2017). “Novel approaches for portfolio construction using second order stochastic dominance”. In: *Computational Management Science* 14(2), pp. 257–280.

Von Neumann, John, Oskar Morgenstern, and Harold William Kuhn (2007). *Theory of games and economic behavior (commemorative edition)*. Princeton university press.

Whang, Yoon-Jae (2019). *Econometric Analysis of Stochastic Dominance: Concepts, Methods, Tools, and Applications*. Cambridge University Press.

Whitmore, George A (1970). “Third-degree stochastic dominance”. In: *The American Economic Review* 60(3), pp. 457–459.

Wu, G. and Z. Xiao (2002). “An analysis of risk measures”. In: *Journal of Risk* 4, pp. 53–76.