UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

CORSO DI LAUREA MAGISTRALE / SPECIALISTICA IN ECONOMICS AND FINANCE – BANKING AND FINANCE

TESI DI LAUREA

"QUANTILE REGRESSION METHODS IN FINANCE: THE CAViaR CASE"

RELATORE:

CH.MO PROF. MASSIMILIANO CAPORIN

LAUREANDO: ALESSANDRO PARMEGGIANI

MATRICOLA N. 1082404

ANNO ACCADEMICO 2015 – 2016

UNIVERSITA' DEGLI STUDI DI PADOVA

DIPARTIMENTO DI SCIENZE ECONOMICHE ED AZIENDALI "M.FANNO"

CORSO DI LAUREA MAGISTRALE / SPECIALISTICA IN ECONOMICS AND FINANCE – BANKING AND FINANCE

TESI DI LAUREA

"QUANTILE REGRESSION METHODS IN FINANCE: THE CAViaR CASE"

RELATORE:

CH.MO PROF. MASSIMILIANO CAPORIN

LAUREANDO: ALESSANDRO PARMEGGIANI

MATRICOLA N. 1082404

ANNO ACCADEMICO 2015 – 2016

Il candidato dichiara che il presente lavoro è originale e non è già stato sottoposto, in tutto o in parte, per il conseguimento di un titolo accademico in altre Università italiane o straniere.

Il candidato dichiara altresì che tutti i materiali utilizzati durante la preparazione dell'elaborato sono stati indicati nel testo e nella sezione "Riferimenti bibliografici" e che le eventuali citazioni testuali sono individuabili attraverso l'esplicito richiamo alla pubblicazione originale.

Firma dello studente

_____

Alla mia amatissima Diana

RINGRAZIAMENTI

Se sono giunto fino a questo grande traguardo universitario è perché non sono mai stato solo durante il mio cammino.
Per tale motivo sento il dovere ed ho il piacere di fare alcuni brevi ringraziamenti a tutte le persone che mi sono state accanto.
Innanzitutto, mi sento di ringraziare il Prof. Massimiliano Caporin, per avermi seguito con pazienza nelle stesura di questo lavoro.
Un ringraziamento particolare va al Dott. Giovanni Bonaccolto, per la sua disponibilità, per tutto il tempo dedicatomi e per i suoi preziosi consigli.
Voglio sinceramente ringraziare i miei genitori, i miei fratelli, i miei parenti e i miei più cari amici per avermi sostenuto nei periodi più difficili di questi anni, mostrandomi il lato migliore della vita.
Un grazie particolare va infine anche ai miei colleghi di facoltà che mi hanno sempre apprezzato per quello che sono e spinto a dare il meglio.

CONTENTS

INTRODUCTION

Value at risk (VaR) is a risk measure "used by financial institution and their regulators since it was first promoted by J.P. Morgan and RiskMetrics and subsequently adopted in the Basel Accords, beginning in 1988" (Allen and Singh (2010)).

Essentially, it is a number that indicates how much a given portfolio can lose within a certain time period and for a certain confidence level; even if VaR seems an easy concept, its calculation represents an important statistical issue.

The relevance of this problem is mainly due to the usage of the VaR as a tool for understanding the market risk hold by a financial firm/institution. Thus if it is not properly estimated, it can lead to an allocation of capital below or above the optimum, affecting the profitability and/or the financial stability of the firm/institution that uses it.

Therefore, the calculation of the VaR is an issue of great importance that requires "accurate knowledge of the distribution of extreme events. This is a difficult task since the distribution of portfolio returns is not constant over time" (Kouretas, Zarangas (2005)).

During the last years, several approaches have been proposed for modelling the VaR, but most of them have focused on creating the whole distribution of the returns, relying on rigid assumptions of normality or independent and identically distributed returns. An alternative approach was introduced by Engle and Manganelli (2002), who used the quantile regression methodology to overcome the above strict assumptions.

In fact, the advantages of quantile regression are essentially three: "First, regression quantile estimates are known to be robust to outliers, a desirable feature in general and for applications to financial data in particular. Second, regression quantile is a semi-parametric technique and as such imposes minimal distributional assumptions on the underlying data generating process (DGP). Third, our multivariate framework allows researchers to directly measure the tail dependence among the random variables of interest, rather than recovering it indirectly via models of time-varying first and second moments" (White, Kim, Manganelli (2015)).

Thanks to this tool, Engle and Manganelli (2002) created a VaR measure called Conditional Autoregressive Value at Risk (CAViaR), which permits to model the evolution of the quantiles over time directly in an autoregressive framework.

The purpose of my thesis is to reproduce/test one CAViaR specification introduced by Engle and Manganelli and compare the results obtained from the estimation of three models made from that CAViaR specification for three comparable portfolios.

The thesis is divided in five chapters: Chapter 1 reviews the literature on regression quantiles, while Chapter 2 introduces the CAViaR models made by Engle and Manganelli (2002). In Chapter 3, I will present some implementations and developments on the original CAViaR, while in Chapter 4 I will introduce my data sets and my models. Chapter 5 reports the empirical results and provides few conclusions.

FIRST CHAPTER: QUANTILE REGRESSION

1.1 Introduction

Quantile regression has born trying to solve an issue related to the "linear regression model and associated estimation method of least squares" (Koenker (2005)).

Indeed, the linear regression curve has a huge disadvantage: it explores just the averages of the distributions related to the set of covariates. The mean is, in fact, only one of the tools statisticians use to investigate the data; other tools (e.g. measures of spread, skewness, histograms, kurtosis, boxplots, etc.) are often used to gain further insight on data's features.

Therefore, the mean gives a partial picture of a single distribution and, consequently, regression curve (that is based on the mean itself) gives an incomplete picture for a set of distributions (Mosteller and Tuckey (1977)).

Koenker tried to overcome this problem proposing a different statistic tool: the Quantile Regression.

1.2 Quantile regression

Linear regression represents the dependent variable as a linear function of one or more independent variables, subject to a random 'error' term; it estimates the mean value of the dependent variable for given levels of the independent variables.

For this type of regression, where we want to understand the central tendency in a dataset, OLS (ordinary least squares) is a very effective method.

The problem arises when we want to go beyond the mean value and/or collecting information about the extremes of a data set by exploring the quantiles; in these cases, OLS loses its efficacy.

It is inside this framework that the quantile and the concept of quantile regression grew.

The quantiles 'objective is to divide a dataset into parts and Quantile Regression tries to extend this idea to the estimation of conditional quantile functions, i.e., as Allen and Singh (2009) said, "models in which quantiles of the conditional distribution of the response variable are expressed as functions of observed covariates".

This was technically achieved by Koenker thanks to a method that can be seen as an extension of classical least squares estimation of conditional mean models: ordinary least squares minimizes the differences between the observed dependent variable and the responses predicted by the linear approximation of the independent variables (i.e. the error). Instead, median regression (i.e. the central special case), also known as least-absolute-deviations (LAD) regression, minimizes a sum of absolute errors. The remaining conditional quantile functions are found by minimizing an asymmetrically weighted sum of absolute errors.

The group of estimated conditional quantile functions offers a more complete view of the effect of the independent variables on the features of the distribution (location, scale, shape, etc.) of the dependent variable (Allen and Singh (2009)).

1.3 Linear regression vs Quantile regression

The purpose of analyzing a data set by a regression is to find out the behavior of a dependent variable given the information contained in a group of explanatory/independent variables.
For doing so, we can use the so-called Ordinary Least Squares approach: it permits us to specify a linear regression model and estimate its unknown parameters by minimizing the sum of its squared errors.

OLS can achieve some good properties called “BLUE” (i.e. it can became the best, linear, and unbiased estimator) if four assumptions are respected:

1. The independent variables are non-stochastic
2. The expectations of the error term are zero
3. Homoscedasticity, i.e. the variance of the error terms is constant
4. No autocorrelation

These assumptions are quite strict and it is very common that data violate some of them, with the consequence that OLS is not the best, linear, unbiased estimator anymore (Montenegro (2001)).

In this case, Quantile Regression can be used to deal with issues arise from the usage of OLS.

For instance, in the very common case in which the error terms are not constant across a distribution, Quantitative Regression (QR) is more robust than OLS.
Indeed, if the errors are highly non-normal distributed, the homoscedasticity assumption of OLS is violated and it is not efficient anymore. On the contrary, QR is more flexible for modeling data with heterogeneous conditional distributions (it is considered a semiparametric model as it avoids assumptions about the parametric distribution of the error process).
Data of this type are common in many fields, including econometrics, survival analysis, and ecology (Koenker and Hallock (2001)).

Another important issue that arises from the usage of OLS regression is that it assumes that the independent variables affect only the location of the conditional distribution of the response.
On the contrary, QR provides a richer description of the data: it allows us to consider the impact of an independent variable on the entire distribution of the dependent one, and not just on its conditional mean.
In fact, as Koenker (2005) said, "different measures of central tendency and statistical dispersion can be useful to obtain a more comprehensive analysis of the relationship between variables".

Another important advantage of quantile regression, with respect to the OLS regression, is that the QR estimates are more robust to extreme outliers.
"In quantile regression, the median estimator minimizes the symmetrically weighted sum of absolute errors (where the weight is equal to 0.5) to estimate the conditional median function, other conditional quantile functions are estimated by minimizing an asymmetrically weighted sum of absolute errors, where the weights are functions of the quantile of interest. This makes quantile regression robust to the presence of outliers" (Allen and Singh (2009)).

Furthermore, QR is invariant to monotonic transformations, so the quantiles of $h(y)$ (a monotone transformation of y) are $h(Q_q(y))$, and the inverse transformation can be used to translate the results back to y.
This is not possible for the mean as $E[h(y)] \neq h[E(y)]$ (Baum (2013)).

1.4 Quantiles via Optimization

Quantile regression is the result of a simple idea: the generalization of the concept of a univariate quantile to a conditional quantile given one or more independent variables (Colin (Lin) Chen).

Therefore, for understanding quantile regression is necessary to introduce and comprehend quantiles. Gilchrist (2001) describes a quantile as "simply the value that corresponds to a specified proportion of an (ordered) sample of a population…Quantiles hereby mark the boundaries of equally sized, consecutive subsets".

From a statistic point of view:

Let X be any real-valued random variable, characterized by its (right-continuous) distribution function

$$F(x) = Prob(X \leq x)$$

We can define the τth quantile of X as the inverse function

$$F^{-1}(x) = inf\{x: F(x) \geq \tau\}$$

where $0 < \tau < 1$.

In particular, the median, $F^{-1}\left(\frac{1}{2}\right)$, plays a key role.

Quantiles are strictly related to the operations of ordering and sorting the sample observations; therefore, it is easy to understand that we can define them through a simple alternative expedient (i.e. an optimization problem). The idea is well expressed by Koenker and Hallock (2001): "just as we can define the sample mean as the solution to the problem of minimizing a sum of squared residuals, we can define the median as the solution to the problem of minimizing a sum of absolute residuals".

In particular, from a technical point of view, for guaranteeing that there are the same number of observations above and below the median, we should have a symmetrical piecewise linear absolute value function (this permits that the minimization of the sum of absolute residuals is equal to the number of positive and negative residuals).

From a mathematical point of view, this can be written as:

For a random sample $\{y_1, y_2, \dots, y_n\}$ of Y, the sample median is

$$min_{\xi \in \mathbb{R}} \sum_{i=1}^{n} |y_i - \xi|$$

Since the symmetry of the absolute value produces the median, thus if we minimize a sum of asymmetrically weighted absolute residuals (i.e. giving different weights to positive and negative residuals) we can get the quantiles.

This can be reached by solving

$$min_{\xi \in \mathbb{R}} \sum_{i=1}^{n} \rho_\tau(y_i - \xi)$$

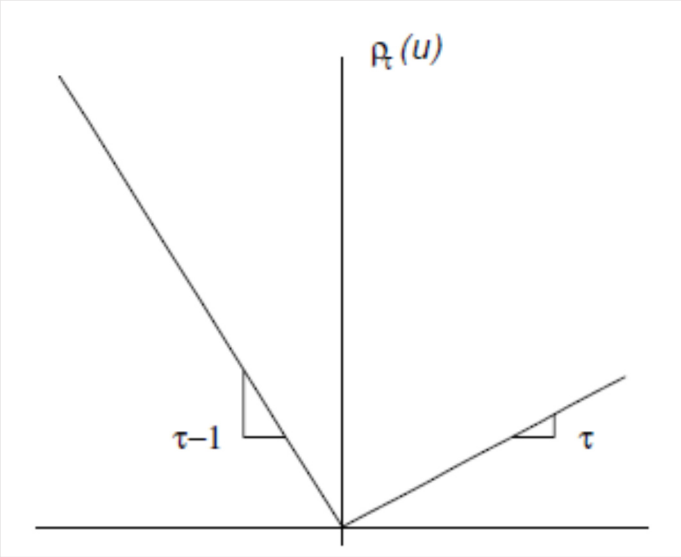
Where $\rho_\tau(\cdot)$ is the tilted absolute value function as shown in Figure 1.

This minimization problem gives the τth sample quantile as its solution.

To see that this is true, it is only necessary to compute the directional derivative of the objective function with respect to ξ, taken from the left and from the right (Allen and Singh (2009)).

For concluding this Section, it is important to highlight the most important thing we have done, i.e. the "the fact that we have expressed the problem of finding the τth sample quantile, a problem that might seem inherently tied to the notation of an ordering of the sample observations, as the solution to a simple optimization problem. In effect, we have replaced *sorting* by *optimizing*. This will prove to be the key idea in generalizing the quantiles to a much richer class of models" (Koenker (2005)).

Figure 1. Quantile regression ρ function



where $\rho_\tau(u) = u(\tau - I(u < 0))$, with $0 < \tau < 1$

1.5 From Quantiles to Quantile Regression

The observation that the unconditional quantiles can be expressed as the solution of an optimization problem, leads us to express conditional quantiles in an analogous way.

QR transforms a conditional distribution function into a conditional quantile function by cutting it into segments. These segments describe, with the use of quantiles, the cumulative distribution of a dependent variable conditional to some explanatory variables (wikibooks.org (2014)).

This can be well illustrate by comparing QR with OLS:

For a random sample $\{y_1, y_2, \dots, y_n\}$, we know that the sample mean (i.e. an estimate of the unconditional population mean) solves the problem

$$min_{\mu \in \mathbb{R}} \sum_{i=1}^{n} (y_i - \mu)^2$$

If we now replace the scalar μ by a parametric function $\mu(x, \beta)$ and solve

$$min_{\beta \in \mathbb{R}^p} \sum_{i=1}^{n} (y_i - \mu(x_i, \beta))^2$$

we obtain an estimate of the conditional expectation function $E(Y|x)$.

In quantile regression, we proceed exactly in the same way.

Let focus on the median case:

Since the sample median (i.e. a scalar), solves

$$min_{\xi \in \mathbb{R}} \sum_{i=1}^{n} |y_i - \xi|$$

To obtain an estimate of the conditional median function, we simply replace the scalar ξ in the previous equation by the parametric function $\xi(x_i, \beta)$ and set τ = 0.5.

Now, to obtain the estimates of the other conditional quantile functions we just replace absolute values by $\rho_\tau(\cdot)$, and solve

$$min_{\xi \in \mathbb{R}^p} \sum_{i=1}^{n} \rho_\tau(y_i - \xi(x_i, \beta))$$

It should be noticed that here, as opposed to OLS, the minimization is done for each subgroups defined by ρ_τ where the estimation of the τth quantile function is reached with the parametric function $\xi(x_i, \beta)$.

The minimization problem, when $\xi(x_i, \beta)$ is expressed as a linear function of parameters, can be solved very efficiently by linear programming methods (Koenker and Hallock (2001)).

1.6 Linear Quantile Regression

Let $A = (x_1, \dots, x_n)$ denote the matrix of n observed vectors of the random vector X, and let $y = (y_1, \dots, y_n)$ denote the n observed responses. The model for linear quantile regression is

$$y = A'\beta + \epsilon$$

where $\beta = (\beta_1, \dots, \beta_p)'$ is the unknown p-dimensional vector of parameters and $\epsilon = (\epsilon_1, \dots, \epsilon_n)'$ is the n-dimensional vector of unknown errors.

As was described in previous section, the τth quantile regression estimator is a solution of

$$min_{\beta \in \mathbb{R}^p}[\sum_{i \in \{i: y_i \geq x_i'\beta\}} \tau |y_i - x_i'\beta| + \sum_{i \in \{i: y_i < x_i'\beta\}} (1-\tau)|y_i - x_i'\beta|]$$

The τth conditional quantile function is just the linear function composed by the τth regression quantile $\beta(\tau)$ and the inverse of the matrix consisting of n observed vectors of the random vector, i.e.

$$Q_y(\tau|A) = A'\beta(\tau)$$

What about the computational aspects?

The minimization problem seen above can be reformulated as a linear programming problem

$$min_{\beta^+, \beta^-, u^+, u^- \in \mathbb{R}^{2k} \times \mathbb{R}_+^{2n}} \{\tau 1_n' u^+ + (1-\tau) 1_n' u^- | X(\beta^+ - \beta^-) + u^+ - u^- = Y\}$$

where

$\beta_j^+ = \max(\beta_j, 0), \beta_j^- = -\min(\beta_j, 0)$

$u_j^+ = \max(u_j, 0), u_j^- = -\min(u_j, 0)$

This linear programming problem can be efficiently solved using some specific simplex algorithm.

For instance, the median regression algorithm of Barrodale and Roberts (1974) was largely used and can be easily modified for general quantile regression problems.

The way the algorithm works was well described by Koenker and Hallock (2000): "The Barrodale and Roberts approach typifies the class of exterior point algorithms for solving linear programming problems: we travel from vertex to vertex along the edges of the polyhedral constraint set, choosing at each vertex the path of steepest descent, until we arrive at the optimum".

Theoretically, the number of iterations can grow very fast with the sample size; nevertheless, this algorithm is still very used when the data set contains less than tens of thousands of observations (Colin (Lin) Chen).

For solving the problems arising from large data sets, some alternative approaches have been developed; among them, we can find the interior point approach of Karmarkar (1984). The way it works was described by Colin (Lin) Chen: it "solves a sequence of quadratic problems in which the relevant interior of the constraint set is approximated by an ellipsoid: instead of traversing the outer surface, it takes Newton steps from the interior of a deformed version of the constraint set toward the boundary".

In addition to this algorithm, other approaches can work quite well when we use a huge data set; in fact, "Portnoy and Koenker (1997) have shown that a combination of interior point methods and effective problem preprocessing render large scale quantile regression computation competitive with least squares computations for problems of comparable size" (Colin (Lin) Chen).

Other approaches, besides the interior point method, have been provided for solving the linear programming problems. One of these is the finite smoothing algorithm of Madsen and Nielsen (1993): it approximates the objective function "with a smoothing function, so that the Newton-Ralphon algorithm can be used iteratively to obtain the solution after a finite number of loops…The smoothing algorithm extends naturally to general quantile regression" (Colin (Lin) Chen).

1.7 Properties of Quantile Regression

Quantile regression estimator has some important properties that can help to better understand regression results.

Several of these properties can be grouped together and identified by a specific name (i.e. equivariance).

If $\hat{\beta}(\tau; y, X)$ denotes a τth regression quantile based on observations (y, X), the four basic *equivariance* properties of $\hat{\beta}(\tau; y, X)$ are the followings:

- Scale equivariance
  For any $a > 0$ and any $0 < \tau < 1$

$$\hat{\beta}(\tau; aY, X) = a\hat{\beta}(\tau; Y, X)$$
$$\hat{\beta}(\tau; -aY, X) = -a\hat{\beta}(1 - \tau; Y, X)$$

- Shift equivariance
  For any $\gamma \in \mathbb{R}^k$ and any $0 < \tau < 1$

$$\hat{\beta}(\tau; Y + X\gamma, X) = \hat{\beta}(\tau; Y, X) + \gamma$$

- Equivariance to reparametrization of design
  Let A any $p \times p$ nonsingular matrix and $0 < \tau < 1$

$$\hat{\beta}(\tau; Y, XA) = A^{-1}\hat{\beta}(\tau; Y, X)$$

It is important to highlight that also the least-squares estimators share these three properties, even if this is not commonly true for other regression estimators.

In addition to the previous properties, Quantiles enjoy another *equivariance* property:

- Invariance to monotone transformations

Let $h(\cdot)$ be a nondecreasing function on $\mathbb{R}$. Then for any random variable Y

$$h(Q_{Y|X}(\tau)) \equiv Q_{h(Y)|X}(\tau)$$

that is, the quantiles of the transformed random variable h(Y) are simply the transformed quantiles of the original Y.

Example: let $Y = \ln(W)$ and $Q_{Y|X}(\tau) = X\beta(\tau)$, then $Q_{W|X}(\tau) = \exp(X\beta(\tau))$)

This property is much stronger than the other properties; in fact, the mean does not share it:

$$Eh(Y) \neq h(E(Y))$$

With the exception of particular cases, i.e. for affine h or other exceptional circumstances (Koenker (2005)).

Besides the equivariance properties there are also some asymptotic properties:

- Asymptotics for the univariate sample quantile
  Let $\{y_1, y_2, \dots, y_n\}$ be an i.i.d. random sample with distribution F, and

$$\widehat{\xi\tau} = argmin_{\xi} \sum_i \rho_\tau(y_i - \xi)$$

  Then

$$\sqrt{n}\left(\widehat{\xi\tau} - \xi\tau\right) \sim \mathcal{N}\left(0, \frac{\tau(1-\tau)}{f^2\left(F^{-1}(\tau)\right)}\right)$$

  where $f = F'$

- Asymptotics for linear models
  $y_i = x_i^\top \beta_\tau + e_i$, where $e_i \sim F_i$ and $F_i^{-1}(\tau) = 0$
  then the join distribution

$$\sqrt{n}\, V_n^{-\frac{1}{2}}\left(\hat{\beta}_{\tau k} - \beta_{\tau k}\right)_{k=1}^{m} \sim \mathcal{N}(0, I_n)$$

where

$$H_n(\tau) = n^{-1}\sum x_i x_i^{\top} \mathfrak{f}_i(0) \text{ and } \mathrm{D_n} = n^{-1}\sum x_i x_i^{\top};$$

$$V_n = [(\tau_k \wedge \tau_l - \tau_k\tau_l)H_n(\tau_k)^{-1} D_n H_n(\tau_l)^{-1}]_{k,l=1}^{m}$$

Based on the asymptotic normality of the estimators we can find the direct estimation of their variance-covariance matrix:

In an i.i.d. error model,

$y_i = x_i^{\top}\beta + e_i, \qquad e_i \sim F_i$

The variance covariance matrix of $\hat{\beta}_\tau$ is

$$\left(\frac{\tau(1-\tau)}{\mathfrak{f}\left(F^{-1}(\tau)\right)^2}\right)(X^{\top}X)^{-1}$$

where $f\left(F^{-1}(\tau)\right)$ is the common error density evaluated at $F^{-1}(\tau)$

In a non i.i.d. error model,

$y_i = x_i^{\top}\beta + e_i, \qquad e_i \sim F_i$

The variance covariance matrix of $\hat{\beta}_\tau$ is

$$(\tau(1-\tau))((X^{\top}F\,X)^{-1}((X^{\top}X)(X^{\top}FX)^{-1}$$

where

$$F = diag\{f\left(F_1^{-1}(\tau)\right), f_2\left(F_2^{-1}(\tau)\right), \dots, f_n(F_n^{-1}(\tau))\}$$

and f_i is the density function of $e_i$evaluated at its τth quantile $F_i^{-1}(\tau)$

Direct estimation of the asymptotic variance-covariance matrix is not always satisfactory. For this reason, inference for quantile regression parameters can be made with other methods, as regression rank-score tests (it avoids direct estimation of the error densities) and bootstrap methods (it avoids direct estimation of variance-covariance matrix) (He, Wei (2005)).

1.8 Strengths and weaknesses of Quantile Regression

To give an objective overview of the most important features of QR, we can critically summarize them into a short outline:

- **Robustness**. From the outset, an important feature of QR has been its robustness to outliers in the response variable.
  While mean regression tends to follow a single outlier in a close way, the influence of an outlying observation on $\hat{\beta}_\tau$ is limited. Indeed, if we move observations away from the QR, they have no effect in the y-direction whatsoever on the fit. This insensitivity is intrinsic to the nature of the quantiles.
  Nevertheless, we should highlight that outliers are still influential in QR, but the problem is not as severe as in the OLS case.

- **Heteroscedasticity**. QR is a more flexible method with respect to the OLS one, this is because it is a semi-parametric technique and, consequently, it "imposes minimal distributional assumptions on the underlying data generating process" (White, Kim, Manganelli (2015)). This feature permits QR to be superior to OLS, which need to impose strict assumption on the data to give good results.

- **Endogeneity**. In 2000, Abadie, Angrist, and Imbens proposed a weighted quantile regression approach to estimate endogenous treatment effects in observational studies (Koenker, Hallock (2000)). After that, several studies have been conducted to find out causality through the usage of QR.

SECOND CHAPTER: CONDITIONAL AUTOREGRESSIVE VALUE AT RISK

2.1 Introduction

The importance of effective risk management has never been greater: the last financial disasters, happened since the early 90s, have emphasized the need for accurate risk measures for financial institutions.
The financial crisis of 1997-1998 as well as the bankruptcy of several financial institutions have led to the increased price volatility and financial uncertainty.
Such financial uncertainty have increased the possibility of financial institutions to suffer important losses as a result of their exposure to unpredictable market changes.
These events have made investors to become more suspicious in their investment decisions while it has also led for the increased need for a more careful study of price volatility in financial markets.
A direct consequence of this is an intensive research from academics, financial institutions and regulators of the banking and financial sectors to better understand the operation of capital markets and develop sophisticated models to analyze market risk.

The market risk is one of the four types of risk that financial institutions can expose themselves (the other three are the credit, the liquidity and the operational risk).
It is considered as the most significant one since it represents the possible economic loss caused by unpredictable movements in market prices, interest rates, exchange rates and in the volatility of options (causing consequently the reduction in the market value of a portfolio).

The existence of market risk and the financial disasters of last decades have raised the need for the development of practical risk management tools for financial institutions. This need has been reinforced by the Basel Committee of Banking Supervision (1996), that has called for the use of internal market risk management to capital requirement by the financial institutions such as banks and investment firms (Kouretas, Zarangas (2005)).
"As the nature of the risks has changed over time, methods of measuring these risks must adapt to recent experience. The use of quantitative risk measures has become an essential management tool to be placed in parallel with models of returns. These measures are used for investment decisions, supervisory decisions, risk capital allocation, and external regulation. In

the fast-paced financial world, effective risk measures must be as responsive to news as are other forecasts and must be easy to grasp in even complex situations" (Engle, Manganelli (2004)).

In this framework, Value at Risk (VaR) has gained, throughout the years, a lot of popularity, until it became the standard measure of market risk used by financial institutions and their regulators.
The main reason why this instrument has achieved a great success is "essentially due to its conceptual simplicity: VaR reduces the (market) risk associated with any portfolio to just one monetary amount" (Engle, Manganelli (2004)).
In practice, VaR is an estimate of the maximum loss a certain portfolio can suffer under normal market conditions over a specified period of time and with a specified level of confidence.
"The confidence level represents 'extreme market conditions' with a probability that is usually taken to be 99% or 95%. This implies that in only 1% (5%) of the cases [we] will lose more than the reported VaR of a specific portfolio" (Kouretas, Zarangas (2005)).

The use of a single number to summarize many complex bad outcomes is essentially a compromise between the needs of different users and, during the years, this compromise has received the approval of many operators and regulators (Engle, Manganelli (2002)). "VaR is probably the most used measure of risk since the 1996 amendment to the Basel Capital Accord, which proposed that commercial banks with significant trade activity could use their own VaR measure to define how much capital they should set aside to cover their market risk exposure" (Allen, Singh (2010)).

Beside the advantages, the use of VaR brings also some weaknesses:

- This measure is just a numerical indication of the maximum potential amount of losses; consequently, it gives us just a partial picture of the situation. A way to deal with this problem is to use VaR side by side with other financial tools (e.g. CVaR, scenario analysis, stress tests, etc.);
- There is no utility function associated with this measure;
- VaR assumes that assets can be sold at their market price without taking into consideration the liquidity problem;
- This measure is not considered a coherent risk measure. Indeed, a risk measure is coherent if it satisfies four main features: "First. It should not exceed the maximum

possible loss to occur. Second, the proposed risk measure should be greater than the mean loss implying capital adequacy to cover losses. Third, in the event that there is a proportional change in the loss then we require that the risk measure change proportionally as well. Finally, it must satisfy the property of superadditivity, implying that the risk measure calculated for two separate losses should be equal to the risk measure calculated on the sum of the two portfolios" (Kouretas, Zarangas (2005)).

VaR methodology does not satisfy the properties of subadditivity and excess of the mean loss. Given these issues regarding the use of VaR, several alternative risk measures have been developed throughout the years (among them, we can find CVaR, which satisfies all the coherency criteria).

Even if VaR suffers from these weaknesses, it is a measure of risk still widely used, because of its simplicity that permits it to be easily understood.
Moreover, as we said above, it is not useless as a financial tool: if we use it side by side with other financial instruments, VaR can give us a more complete analysis of our portfolio, creating a number that can help us to better understand the entire financial situation we hold (and act consequently).

Even if VaR is theoretically simple, its measurement is a huge statistic challenge and only few methodologies developed so far gives satisfactory solutions (Engle, Manganelli (2002). Indeed, the calculation of VaR requires precise knowledge of the distribution of extreme events, task that results difficult since the distribution of portfolio returns is not constant over time. In addition, we must find an appropriate model for time varying conditional quantiles, since VaR appears to be "nothing more than a specific quantile of future portfolio values subject to current information" (Kouretas, Zarangas (2005)).

As was said by Engle and Manganelli (2002), the main statistic problem consists in forecasting "a value that in each period will be exceeded with probability (1-θ) by the current portfolio, where $\theta \in (0,1)$ represents the confidence level associated to the VaR".
This problem can be summarized in the following statistic problem:

If we denote the time series of portfolio returns by $\{y_t\}_{t=1}^{T}$ and the sample size by T, we want to find VaR_t such that $Pr[y_t < -VaR_t \,|\Omega_t\,] = \theta$, where Ω_t denotes the information set at the end of time t-1.

For solving this problem, we need a methodology that must be reasonable and to be like this, a methodology should solve the following three issues: "1) Provide a formula for calculating VaR_t as a function of variables known at time t-1 and a set of parameters that need to be estimated; 2) Provide a procedure (namely, a loss function and a suitable optimization algorithm) to estimate the set of unknown parameters; 3) Provide a test to establish the quality of the estimate" (Engle, Manganelli (2002)).

Consequently, the central problem to find an appropriate model for time varying conditional quantiles is strictly connected to the issue of finding accurate estimates of the chosen distribution of portfolio returns. In fact, as Engle and Manganelli (2004) argued, if we do not properly estimate the underlying market risk then we cannot reach an efficient allocation of the capital, with the consequence of a reduction of the profitability and/or the financial stability of banks or investment firms, that use this incorrect methodology.

During the years, a lot of methodologies have been developed to estimate the distribution of portfolio returns, but, as was said by Kouretas and Zarangas (2005), "these alternative methodologies have mainly focused on modeling the entire distribution of returns and they are based on the strict assumptions of normality or i.i.d. (independent and identically distributed) returns".
Engle and Manganelli (2004) have proposed an innovative approach that does not model the entire distribution: it just focuses on the regression quantile, which does not entail the strict assumptions seen above.
This innovative methodology, called CAViaR (Conditional Autoregressive Value at Risk), "uses an autoregressive process in order to model the evolution of the regression quantile over time" and "the estimation of the unknown parameters is done with the use of the framework suggested by Koenker and Bassett (1978)" (Kouretas and Zarangas (2005)).
In addition, as was demonstrated by Engle Manganelli (2002), CAViaR estimators are asymptotically efficient and consistent.

2.2 Value at risk models

VaR methodology was created in the 90's with the purpose of giving senior management of financial institutions a single number that could quickly and easily offer a summary about the risk of a portfolio.

Nowadays, it is largely used: in fact, it helps management to better estimate the cost of positions in terms of risk, allowing to allocate risks more efficiently (Engle, Manganelli (2004)).
In addition, VaR is an instrument used by regulators to require financial institutions such as banks and investment firms to meet capital requirements to cover the market risks that they incur because of their normal operations (we can still find it in Basel III in the form of "stressed VaR").
"However, if the underlying risk is not properly estimated, these requirements may lead financial institutions to overestimate (or underestimate) their market risks and consequently to maintain excessively high (low) capital levels. The result is an inefficient allocation of financial resources that ultimately could induce firms to move their activities into jurisdictions with less restrictive financial regulations" (Engle, Manganelli (2004)).
Another probable consequence of the inefficient allocation of capitals is a financial disaster: the best example of this is the 2008 credit crisis, that was the result of the huge underestimation of the risk from toxic mortgage products and of the permission for banks to enjoy excessive levels of leverage on their trading positions (Financial Times, 2012).

During the 90s, many alternative modeling methodologies were proposed to estimate the VaR. The motivation behind that must be found in the characteristics of financial data, which have been firstly documented by Mandelbrot (1963) and Fama (1965). To give a grand summary, these characteristics suggest that the returns of financial assets have leptokurtic distributions, their distributions are negatively skewed and, finally, they exhibit volatility clustering.

The existing models for calculating VaR differ in many aspects, but, as Engle and Manganelli (2001, 2004) pointed out, these alternative methodologies adopted a common general structure, which can be summarized in three points:"1) the portfolio is marked-to-market on a daily basis; 2) the distribution of the portfolio returns is estimated; 3) the VaR of the portfolio is computed".
The main difference among the alternative methodologies is linked to the estimation of the appropriate distribution of the portfolio returns (i.e. the second point).
For a better understanding of this point, we present a classification of VaR methodologies.

At first, these methodologies can be ordered into two broad categories: a) factor models such as RiskMetrics (1996), b) portfolio models such as historical quantiles.
"In the first case, the universe of assets is projected onto a limited number of factors whose volatilities and correlations have been forecast. Thus, time variation in the risk of a portfolio is associated with time variation in the volatility or correlation of the factors. The VaR is assumed

to be proportional to the computed standard deviation of the portfolio, often assuming normality." (Engle, Manganelli (2004)).
In the second case, the portfolio models create historical returns that reproduce the past performance of the current portfolio; then these historical returns are used to build the current VaR by using a statistical model; consequently, the relation between changes in the risk of a particular portfolio and its historical experience is very tight.
Even if some issues may arise in the construction of the historical returns, the most interesting modeling problem concerns how to forecast the quantiles – in fact, VaR is nothing more than a quantile of a distribution of returns (Engle, Manganelli (2002)).
Many different approaches have been employed and Engle and Manganelli (2004) gave a general description of them: "some first estimate the volatility of the portfolio, perhaps by a generalized autoregressive conditional heteroscedasticity (GARCH) or exponential smoothing, and then compute VaR from this, often assuming normality. Others use rolling historical quantiles under the assumption that any return in a particular period is equally likely. A third approach appeals to extreme value theory".

From this very general description, we can start to focus on each of these approaches.

The first class of approaches are fully parametric and includes models such J.P. Morgan's Riskmetrics (1996) and GARCH models. "These procedures combine an econometric model with the assumption of conditional normality for the returns series. Specifically, these models rely on the specification of the variance equation of the portfolio returns and the assumption that the standardized errors are i.i.d. Additionally, when the GARCH methodology is applied we are also required to specify the distribution of the errors, which is usually taken to be the normal one, while it is assumed that the negative returns follow the same process like the rest of portfolio returns" (Kouretas and Zarangas (2005)).

The second approach for estimating the distribution returns is the non-parametric historical simulation. The main features of these approaches are that they do not make any assumption about the distribution of the portfolio returns and that they estimate the VaR as the quantile of the empirical distribution of historical returns from a moving window.
The idea at the base of these methodologies is to select a window, drawn from recent periods (usually anywhere between 6 months to 2 years), and assume that any portfolio return can occur with the same probability. In particular, all the returns that fall outside the selected window have zero probability to occur (Kouretas, Zarangas (2005)).

The so-called hybrid approach (developed by Boudoukh et al. (1998)) falls within this group of VaR models: it combines the historical simulation and Riskmetrics (i.e. the volatility method). "This methodology applies weights to past portfolio returns that decline exponentially" (Kouretas, Zarangas (2005)).

The semiparametric models create the final group.
The advantage of this approach is that it is based on rigorous statistical theory: in fact, as was said by Kouretas and Zarangas (2005), it "offers a parametric form for the tail of a distribution". In particular "This approach focuses on the asymptotic form of the tail, rather than modeling the complete distribution of portfolio returns and therefore we are able to obtain more efficient forecasts of the risk associated with a particular market position" (Kouretas, Zarangas (2005)).

Obviously, each of these models bring some weaknesses:

1) The parametric methodologies have a tendency to assume that the negative extremes follow the same process as the rest of the returns and to create coefficients that underestimate the VaR (this happens because these methodologies fail to take into account the characteristic that the distribution of the portfolio returns have heavy tails). "This underestimation of the VaR as well as possible misspecifications with respect to the variance equation along with the distribution of errors can be corrected by allowing alternative distributions of the errors such as the Gaussian, Student's t and Generalized Error Distribution. However, it is…shown that the GARCH-type models provide satisfactory estimates of the quantile only when a bad event has already occurred" (Kouretas, Zarangas (2005)).

2) The non-parametric historical simulation methodology has several problems: it assumes that for a certain time window (e.g., one year) any return has the same probability to happen, but a return older than a year has zero probability of occurring.
   "It is easy to see that the VaR of a portfolio will drop dramatically just one year after a very bad day. Implicit in this methodology is the assumption that the distribution of returns does not vary over time at least within a year" (Engle, Manganelli (2004)).
   Therefore, this methodology is unsuitable to provide extreme quantiles since it is impossible for us to extrapolate beyond past data.
   "The proposed solution to this problem is the increase of the sample of observations but this would lead to estimates of the VaR which are biased downwards (or upwards) since we

have a mixture of periods with low volatility with periods of high volatility" (Kouretas, Zarangas (2005)).

3) Although the hybrid approach improves the previously methodologies, it has some problems: indeed, the selection of the parameters and the calculation of the VaR do not depend on rigorous statistical theory but they seem to be created ad hoc and built on empirical basis (Kouretas, Zarangas (2005)).

4) Even if the semiparametric models are very interesting, they have two main problems. First, these approaches perform well at very low quantiles but they fail to provide accurate estimations of VaR when the levels are not such extreme (e.g. probability levels of 5%).
   Second, and most important, these models are made inside a framework of i.i.d. variables, which is not consistent with the characteristics of most financial dataset. Consequently, this framework creates a situation where the risk of a portfolio cannot vary with the conditioning information set.
   "Recently, McNeil and Frey (2000) suggested fitting a GARCH model to the time series of returns and then applying the extreme value theory to the standardized residuals, which are assumed to be i.i.d. Although it is an improvement over existing applications, this approach still suffers from the same criticism applied to the volatility models" (Engle, Manganelli (2004)).

2.3 CAViaR

Engle and Manganelli (2004) proposed an alternative semiparametric approach for estimating the VaR, the so-called CAViaR (Conditional Autoregressive Value at Risk).

This technique is based on this simple idea: "it is better to model directly the quantile as it evolves through time instead of attempting to model and estimate the entire distribution of portfolio returns. [In fact] modelling the quantile instead of the entire distribution has the main advantage that we are not required to adopt the set of extreme assumptions which are invoked by alternative methodologies, among them normality or that returns are i.i.d." (Kouretas, Zarangas (2005)).

Another simple intuition, at the base of the CAViaR creation, is linked to the characteristics of financial data that have been verified by numerous empirical works; in particular, one of these

characteristics is the volatility clustering of portfolio. This feature of portfolio returns leads to the understanding that the corresponding distributions are autocorrelated.
Consequently, as Engle and Manganelli (2004) said, "the VaR, which is tightly linked to the standard deviation of the distribution, must exhibit a similar behavior".
Therefore, a natural way to formalize this characteristic, as was suggested by Engle and Manganelli (2004), is to "use some type of autoregressive specification". To do so, they proposed a conditional autoregressive quantile specification, which they called Conditional Autoregressive Value at Risk (CAViaR).

Following Engle and Manganelli (2004), we consider $\{y_t\}_{t=1}^{T}$ as a vector of portfolio returns that is observable. Let denote by θ the probability related to VaR, by x_t a vector of observable variables at time t and by β_τ a p-vector of unknown parameters. We also define $f_t(\beta) = f(x_{t-1}, \beta_\theta)$ to be the θ - quantile of the distribution of the portfolio returns at time t which has been formed at time $t-1$, where we suppressed the θ subscript from β_θ for notational convenience
Therefore, a general formulation of CAViaR can be written as follows:

$$f_t(\beta) = \gamma_0 + \sum_{i=1}^{q} \gamma_i f_{t-i}(\beta) + \sum_{i=1}^{p} a_i l(x_{t-i}, \varphi)$$

where $\beta' = (\alpha', \gamma', \varphi')$ and l is a function of a finite number of lagged values of observables.
Moreover, Engle and Manganelli use the autoregressive terms $\gamma_i f_{t-i}(\beta), i = 1, \dots, q$ for ensuring that the quantile changes "smoothly" over time.
Finally, they use the term $l(x_{t-i}, \varphi)$ to provide a relationship between the θ - quantile $f_t(\beta)$ and the observable variables, which are included in the information set.
We could consider the lagged portfolio returns as the natural choice for x_{t-1}. Indeed as Engle and Manganelli (2004) pointed out "we would expect the VaR to increase as y_{t-1} becomes very negative, as one bad day makes the probability of the next somewhat greater. It might be that very good days also increase VaR as would be the case for volatility models. Hence VaR could depend symmetrically upon $|y_{t-1}|$."
In other words, as y_{t-1} becomes negative then we should expect the VaR to increase while it should decline in good days. Thus, we expect that changes in y_{t-1} will affect symmetrically the VaR (Kouretas, Zarangas (2005)).

From this general setting, we can estimate different models by choosing different specifications for the function l. Engle and Manganelli (2002) proposed four alternative CAViaR specifications.

The first specification is called *Adaptive,* which takes the following formulation:

$$f_t(\beta_1) = f_{t-1}(\beta_1) + \beta_1\{[1 + \exp(G[y_{t-1} - f_{t-1}(\beta_1)])]^{-1} - \theta\}$$

where G is some positive finite number.
As G→ ∞, the last term converges almost certainly to $\beta_1\big[I\big(y_{t-1} \leq f_{t-1}(\beta_1)\big) - \theta\big]$, where $I(\cdot)$ represents the indicator function; for finite G this model is a smoothed version of a step function.

The intuition behind the *adaptive* specification tells us that: "whenever you exceed your VaR you should immediately increase it, but when you don't exceed it, you should decrease it very slightly" (Engle, Manganelli (2004)). Such a strategy can reduce the probability to observe a sequence of hits but, at the same time, it does not bring the number of hits to zero (Kouretas, Zarangas (2005)).

Engle and Manganelli (2002) also pointed out that this CAViaR specification "learns little from returns which are close to the VaR or which are extremely positive. [In fact] it increases the VaR by the same amount regardless of whether the returns exceeded the VaR by a small or a large margin". They also highlighted that this model specification has a unit coefficient on the lagged VaR.

A second specification is called *Symmetric Absolute Value* (SAV) and its formulation is given by:

$$f_t(\beta) = \beta_1 + \beta_2 f_{t-1}(\beta) + \beta_3|y_{t-1}|$$

"This model responds symmetrically to past portfolio returns and it is mean reverting since the coefficient of the lagged VaR is not constrained to equal one. Furthermore, we could properly specify this quantile specification using a GARCH model with the standard deviation (and not

the variance) is considered to follow a symmetric distribution with i.i.d. errors" (Kouretas, Zarangas (2005)).

The *Asymmetric Slope* (AS) is the third model created by a specification of the l function. It is written as follows:

$$f_t(\beta) = \beta_1 + \beta_2 f_{t-1}(\beta) + \beta_3 (y_{t-1})^+ + \beta_4 (y_{t-1})^-$$

where $(x)^+ = \max(x, 0)$, $(x)^- = -\min(x, 0)$

The *Asymmetric Slope* model allows for an asymmetric response to positive and negative past portfolio returns and it is mean reverting.

"As with SAV model, this specification would be correctly specified by a GARCH process with the standard deviation following this time an asymmetric distribution with i.i.d. errors" (Kouretas, Zarangas (2005)).

The last specification is called *Indirect GARCH(1,1)* which is mean reverting and responds symmetrically to past returns (as we saw in the SAV specification).

"This specification can be correctly modeled under the assumption that the underlying data process follows a true GARCH(1,1) with an i.i.d. error distribution" (Kouretas, Zarangas (2005)).

The mathematical expression of this specification is:

$$f_t(\beta) = (\beta_1 + \beta_2 f^2{}_{t-1}(\beta) + \beta_3 y_{t-1}^2)^{1/2}$$

It is interesting to note that CAViaR specifications are more general than the fitted GARCH models, in fact, they can allow for modelling various forms of non-i.i.d. error distributions. "These models can be used for situations with constant volatilities, but changing error distributions, or situations where both error densities and volatilities are changing" (Engle and Manganelli (2004)).

2.4 Estimation of parameters & Variance and Covariance Matrix

The next step to the analysis is the estimation of the parameters of the different CAViaR models.

They are estimated using linear and non-linear quantile techniques that we have already discussed in the first chapter; consequently, we are going to give just the mathematical expressions:

Let consider a sample of observations $y_1, \dots, y_T$ generated by the following model:

$$y_t = x_t'\beta^0 + \varepsilon_{\theta t} \quad Quant_\theta(\varepsilon_{\theta t}|x_t) = 0$$

where x_t is a p - vector of regressors, $Quant_\theta(\varepsilon_{\theta t}|x_t)$ is the θ - quantile of $\varepsilon_{\theta t}$ conditional on x_t and $f_t(\beta) \equiv x_t\beta$
Then the θ^{th} regression quantile is defined as any $\hat{\beta}$ that solves:

$$(1) \qquad min_\beta \frac{1}{T}\left[\theta - I\left(y_t < f_t(\beta)\right)\right][y_t - f_t(\beta)]$$

It is important to remember that within this framework, the only assumption required is the appropriate specification of the quantile process and, in particular, we do not need to specify the entire distribution of the error terms (Kouretas, Zarangas (2005)).

Let now focus on the derivation of the variance-covariance matrix: for estimating it, Engle and Manganelli (2004) considered the case where $\hat{\beta}$ is a non-linear regression quantile estimator and they showed that this estimator is consistent and asymptotically normal. In addition, they proved that there is a consistent estimator of the variance-covariance matrix and they derived its asymptotic distribution (Kouretas, Zarangas (2005)).

Let consider the model:

$$(2) \qquad \begin{aligned} y_t &= f(y_{t-1}, x_{t-1}, \dots, y_1, x_1; \beta^0) + \varepsilon_{t\theta} \qquad Quant_\theta(\varepsilon_{t\theta}|\Omega_t) = 0 \\ &\equiv f_t(\beta^0) + \varepsilon_{t\theta} \qquad\qquad t = 1, \dots, T \end{aligned}$$

where $f_1(\beta^0)$ represents some given initial condition, x_t is a vector of exogenous variables,

$\beta^0 \in \mathbb{R}^p$ is the vector of true unknown parameters and $\Omega_t = [y_{t-1}, x_{t-1}, \dots, y_1, x_1, f_1(\beta^0)]$ is the information set available at time t.
Let $\hat{\beta}$ be the parameter vector that minimizes (1)

Let then consider the following Theorems, created and developed by Engle and Manganelli (2002):

- **Theorem 1 (Consistency)** – In model (2), under C0-C7 in Appendix A,

$$\hat{\beta} \xrightarrow{p} \beta^0$$

  where $\hat{\beta}$ is the solution to (1).

- **Theorem 2 (Asymptotic Normality)** – In model (2), under AN1-AN4 in Appendix A and the conditions of Theorem 1,

$$\sqrt{\frac{T}{\theta(1-\theta)}} A_T^{\frac{-1}{2}} D_T\left(\hat{\beta} - \beta^0\right) \xrightarrow{d} N(0, I)$$

  where $A_T = E[T^{-1}\nabla' f(\beta^0)\nabla f(\beta^0)]$, $D_T = E[T^{-1}\nabla' f(\beta^0)\mathrm{H}\nabla f(\beta^0)]$ and H is a diagonal matrix with typical entry $h_t(0|\Omega_t)$.

- **Theorem 3 (Variance-Covariance Matrix Estimation)** – Under VC1-VC2 in Appendix A and the conditions of Theorem 1 and 2,

$$\hat{A}_T \xrightarrow{p} A_T \text{ and } \widehat{D}_T \xrightarrow{p} D_T$$

  where $\hat{A}_T = T^{-1}\nabla' f(\hat{\beta})\nabla f(\hat{\beta})$
  and $\widehat{D}_T = (2T\hat{c}_T)^{-1} \sum_{t=1}^{T} I(|y_t - f_t(\hat{\beta})| < \hat{c}_T)\nabla' f_t(\hat{\beta})\nabla f_t(\hat{\beta})$

We denote the conditional density of $\varepsilon_{t\theta}$ evaluated at 0 by $h_t(0|\Omega_t)$, the $(1, p)$ gradient of $f_t(\beta)$ by $\nabla f_t(\beta)$ and define $\nabla f(\beta)$ to be a (T, p)matrix with typical row $\nabla f_t(\beta)$.
Appendix A is formed by technical assumptions, which permits these results to hold.

If we want to summarize the results of the theorems, we can say that Theorems 1 and 2 prove that $\hat{\beta}$ (i.e. the non-linear regression quantile estimator) is consistent and asymptotically normal, while Theorem 3 provides a consistent estimator of the variance-covariance matrix.
The proofs are simple extensions of Weiss (1991) and Powell (1984, 1986, and 1991) and are omitted.

"Regarding the variance-covariance matrix, note that $\hat{A}_T$ is simply the outer product of the gradient. Estimation of the $\widehat{D}_T$ matrix is less straightforward, as it involves the term $h_t(0|\Omega_t)$. Following Powell(1984, 1986, 1991), we propose an estimator that combines kernel density estimation with the heteroskedasticity-consistent covariance matrix estimator of White (1980)" (Engle, Manganelli (2002)).

The asymptotic distribution of the estimator $\hat{\beta}$ allowed Engle and Manganelli to conduct hypothesis tests on the quantile models. They also proposed a new test for the evaluation of the alternative specifications, which has better power properties than other existing tests (Kouretas, Zarangas (2005)).

THIRD CHAPTER: IMPLEMENTATION & DEVELOPMENTS

3.1 Introduction

In this section, we will present the way to implement the CAViaR model via statistical software. We will start from the description of Engle and Manganelli's approach and then we will try to translate it into codes; it is important to highlight that we will focus just on one CAViaR specification (the Symmetric Absolute Value one).
We are going to use some empirical data for reproducing the CAViaR model and for understanding if it works properly.
At the end of the chapter, I will present and discuss some improvements on the original model.

3.2 The original model

Engle and Manganelli (2004) gave a description on how to implement the theory they developed on CAViaR by using an empirical experiment.

As was said by Engle and Manganelli (2004), the first step for implementing CAViaR methodology on real data is to "construct the historical series of portfolio returns and to choose a specification of the functional form of the quantile".
For this reason, they took a sample of 3,392 daily prices from Datastream for two companies (General Motors (GM) and IBM) and one index (the S&P 500). Then they computed the daily returns as 100 times the difference of the log of the prices (the samples covered a period of 14 years (from April 7, 1986, to April 7, 1999)).
Then, they split the sample into two parts: the first 2,892 observations and the last 500 observations; this was done for making some final tests.

To give a comprehensive example of the theory and to understand which model was more appropriate for the data sets, Engle and Manganelli (2004), estimated the 1% and the 5% 1-day VaRs, using all the four CAViaR specifications described in Section 2.3.
In particular, they highlighted how all the four models considered satisfy assumptions C1, C7 and AN1 of Appendix A and how all the models are both continuous and continuously differentiable in β (the other assumptions in Appendix A are technical assumptions that are impossible to verify in finite samples).

For computing the VaR series with the CAViaR models, it is important to set an optimization routine.
This routine should be initialize in its first step/quantile (i.e. $f_1(\beta)$) and Engle and Manganelli chose to do it by using the empirical θ-quantile of the first 300 observations.

For making all the computation, the two researchers used as software MATLAB 6.1 and as optimization routines the Nelder–Mead simplex algorithm and a quasi-Newton method; Engle and Manganelli chose to use the functions *fminsearch* and *fminunc* as optimization algorithms and the loops to compute the recursive quantile functions were coded in C.

What is very interesting in the computational part is the procedure they followed for optimizing the model:
Engle and Manganelli started by trying to find out the best initial values to feed the optimization algorithm; for doing so, they generated *n* vectors using a uniform random number generator between zero and one.
Then, they computed the following regression quantile (RQ) function

$$min_{\beta} \frac{1}{T} \left[\theta - I\left(y_t < f_t(\beta)\right)\right][y_t - f_t(\beta)]$$

for each of these *n* vectors.

After this step, they selected the *m* vectors, which produced the lowest RQ criterion, as initial values for the optimization routine.
The researchers set, arbitrary, $n = [10^4, 10^5, 10^4, 10^4]$ and $m = [10, 15, 10, 5]$ respectively for, the symmetric absolute value, the asymmetric slope, the Indirect GARCH, and the adaptive models.

The second step for optimizing the model was to run the simplex algorithm for each of these *m* initial values. The two researchers then fed the optimal parameters found (i.e. the β's) to the quasi-Newton algorithm and chose the new optimal parameters as the new initial conditions for the simplex.
For finding the best values (i.e. to reach the global minimum instead of a simple local minimum), they repeated this procedure until the convergence criterion was satisfied.

In particular, tolerance levels for the function and the parameters values were set to 10^{-10}. Finally, Engle and Manganelli selected the vector that produced the lowest RQ criterion as optimal parameters.

It is important to highlight that there exists alternative optimization routines (as the interior point algorithm for nonlinear regression quantiles, suggested by Koenker and Park (1996)), but I will not treat them because I preferred to focus on Engle and Manganelli's optimization loop, trying to analyze and improve it.

A further natural step is to understand if the parameters are well estimated (i.e. if they are significantly different from zero) and for doing this, we need a statistical test.
For this reason, I chose the so-called t-test, which requires to compute the standard errors and the variance–covariance matrix; for making them, Engle and Manganelli used the formulas described in Theorems 2 and 3 in Section 2.4:

After computed
$\hat{A}_T = T^{-1}\nabla' f(\hat{\beta})\nabla f(\hat{\beta})$ and
$\widehat{D}_T = (2T\hat{c}_T)^{-1}\sum_{t=1}^{T} I(|y_t - f_t(\hat{\beta})| < \hat{c}_T)\nabla' f_t(\hat{\beta})\nabla f_t(\hat{\beta})$

The variance-covariance matrix is:

$$VCmatrix = \theta(1-\theta) * (\widehat{D}_T \backslash \hat{A}_T / \widehat{D}_T)/T$$

Where the formula to compute $\widehat{D}_T$ were implemented using *k*-nearest neighbor estimators, with $k = 40$ for 1% VaR and $k = 60$ for 5% VaR.

3.3 The translation of the model into codes

At this point is important to translate all Engle and Manganelli's work into codes; we will then use these codes as the base for analyzing some specific models and data sets.

As the two researchers said, the first step for implementing the methodology on real data is to choose the data set and create the historical series of portfolio returns; for this reason (and also for make comparisons with the results of the two researchers) I took a sample of 3,581 daily prices of the S&P 500 index from Datastream. The sample covered a period of approximately 15 years (from August 27, 2001, to November 23, 2015).

Then I computed the daily returns as 100 times the difference of the log of the prices (as was done by Engle and Manganelli).

The "second" initial step is to choose a "specification of the functional form of the quantile" (Engle, Manganelli (2004)). My choice focused on the *Symmetric Absolute Value* (SAV) model:

$$f_t(\beta) = \beta_1 + \beta_2 f_{t-1}(\beta) + \beta_3 |y_{t-1}|$$

This model responds symmetrically to past portfolio returns (i.e. positive or negative returns have a symmetric impact on VaR.) and it is mean reverting.

For the remaining computations, I completely followed the two researchers' procedure, except for two things:

1) For initializing the optimization routine, I used the entire sample to compute the empirical θ-quantile (instead of the first 300 observations);
2) All the codes, loops and computations were made using just one software: MATLAB R2014a.

In the following table, we can see the results for the 1% and the 5% VaR (β's and associated p-values)

Table 1. The 1% & 5% original CAViaR estimations for the S&P 500 index

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.0897	3.0627e-04	-0.0311	0.0240
AutoregQuant	0.8976	0	0.8952	2.8539e-246
abs(y)	-0.2451	4.4062e-21	-0.2008	5.2334e-04

The results presented in the Table above share many common characteristics with those presented by Engle and Manganelli (2004) for their US data set (which contained General Motors, IBM and the S&P 500 index). One of these matching characteristics is that the autoregressive term is always very significant, suggesting us, that "the phenomenon of clustering of volatilities is relevant also in the tails" (Engle, Manganelli (2004)).

It is important to highlight that the data set I chose is different with respect to the one picked by Engle and Manganelli: even if we both chose to extrapolate the data set from the same index (i.e. the S&P 500), the years we took are different and so are the results.

In particular, the period I extrapolated takes also into account the years of the global financial crisis started in 2008: this huge volatility period has a strong influence on the results. Nevertheless, the parameters are all statistically significant (with the exception of the constant term in the 5% VaR, which is significant just at the 5% level) and both the autoregressive components and the absolute values have a strong economic impact.

I would conclude that the model is well coded and give me interesting results from both a statistic and an economic point of view.

3.4 Improvements

Analyzing data and reaching good estimation is very important from an economic point of view, but also time is precious: if you reach good estimation in less time you become more efficient and you can take decision faster, that led, at the end, to earn (or not to lose) money.

This is exactly the way through which I tried to improve Engle and Manganelli's codes, i.e. reducing the time of the calculations, while maintaining the quality of the estimations.

For doing so, I experimented four different types of procedure for optimizing the model:

1) I diminished the number of the *n* vectors (i.e. the vectors created by a uniform random number generator between zero and one) used to find out the most suitable initial values to put inside the optimization routine.

I used 10^3 vectors instead of 10^4.

2) I diminished the initial *n* vectors to 10^2 and I fed them directly inside the optimization algorithm. Then, I chose the initial vector that produced the lowest RQ criterion as optimal parameters.
   In particular, I just used the simplex algorithm and I used it once per vector.

3) I diminished the number of the *n* vectors from 10^4 to 10^3. Moreover, I removed from the optimization algorithm the quasi-Newton algorithm.
   Manganelli himself suggested this choice in his own publications ("VAR for VaR: Measuring tail dependence using multivariate regression quantiles", 2015).

4) Instead of using *n* initial vector, I chose to use an algorithm that finds the best initial vector of parameters to feed then into the optimization algorithm. This algorithm is called Genetic Algorithm and it is a heuristic method that can minimize a given function.
   In other words, I used two minimization functions: the first one to find out the best parameters that will be used to initialize the second function.
   The second minimization algorithm is equal to the one described in the third method: it just uses the simplex algorithm (i.e. I dropped out the quasi-Newton one).

The next step is to understand which procedure is the best one.

For this reason, I ran all of them and I analyzed both the precisions of the results and the time they took to complete all their iterations (I used the same data set used in previous Section).

The results, for all the four methods, are the same I showed in Table 1 in previous Section; this does not came as a surprise, because I tried to find out the processes that can produce me, at least, the same quality of results that I found with the original procedure.

The most interesting results are the times the four methods took with respect to the original one:

- the original process took 14 seconds
- the first one took 8 seconds
- the second one took 18 seconds
- the third one took 6 seconds
- the last one took 5 seconds

It is easy to analyze the time results: the best method seems to be the one that uses the Genetic Algorithm to find out the best initial values.

Nevertheless, we can also make some considerations about the other procedures: if we move from the original to the first process, we save more than 40% of the time; this is reasonably due to the reduced number of initial vectors that I used. This strategy appears to be also one of the reason why the third method gives us such a good time-result. In particular, if we compare the third and the original method, we could conclude that by diminishing the number of initial vectors and by eliminating the quasi-Newton algorithm, we can save a lot of time (more than 57%), without diminishing the quality of the results.

The second procedure appears to be the worst one; the most likely reason could be the usage of 100 vectors that I fed directly to the optimization loop and, even if the loop is less heavy than the one used in the original process, this requires a huge amount of time.

It is important to highlight that all the consideration made above are restricted to the case of a model composed by three parameters. Nevertheless, we can learn something from this experiment and start to orientate our decision to use, in future analysis, some procedures than others.

In this section, we introduced a new kind of algorithm (Genetic Algorithm), which can give us some advantages in saving time during the computational process. For this reason, it is important to know something more about it.

3.5 The Genetic Algorithm

“The genetic algorithm is a method for solving both constrained and unconstrained optimization problems that is based on natural selection, the process that drives biological evolution. The genetic algorithm repeatedly modifies a population of individual solutions. At each step, the genetic algorithm selects individuals at random from the current population to be parents and uses them to produce the children for the next generation. Over successive generations, the population "evolves" toward an optimal solution.” (The MathWorks, Inc. (2014)).

This algorithm can solve different optimization problems that are not well suited for standard optimization algorithms, comprising problems where the objective function is discontinuous, nondifferentiable, stochastic, or highly nonlinear. (The MathWorks, Inc. (2014))

To give an overview on how the algorithm works, it is necessary to introduce some specific terminology and basic concepts:

- The *fitness function* is the function we want to optimize (in standard optimization algorithms it is usually called objective function).

- The *fitness value* of an individual is the value that the fitness function receives for that individual (it is also called score of the individual)

- An *individual* is any point to which we can implement the fitness function. Sometimes, an individual can be called *genome* and the vector that entries an individual is called *genes*.

- A *population* is a selection of individuals. At each iteration, the genetic algorithm makes some computational steps on the current population to produce a new one. Each following population is called a new *generation*.

- *Diversity* denotes the average distance between individuals in a population. It is a crucial element for the algorithm; in fact, it permits the algorithm to search for a larger region of the space to find new individuals.

- The genetic algorithm creates the next generation by choosing certain individuals in the current population (called *parents*), and uses them to create individuals in the next generation (called *childrens*). It is more likely that the algorithm chooses parents with a better fitness values.

- "The selection function chooses parents for the next generation based on their scaled values from the fitness scaling function" (The MathWorks, Inc. (2014)). The algorithm can choose an individual more than once for becoming a parent; in this case, it gives its genes to more than one child.

(The MathWorks, Inc. (2014))

Let now focus on the steps the algorithm takes when it runs (let remember that all the consideration we are going to make are restricted to the Genetic Algorithm created by Matlab developers):

1. The algorithm starts creating a random initial population.

2. The algorithm creates a series of new populations.
   In every step, the algorithm creates the new generation by using individuals in the current generation.
   To create the new generation, the algorithm performs the following steps:

   a. Scores each member of the current population by computing its fitness value (i.e. the value of the target function - in our case the RQ function)
   b. Scales the raw fitness scores to convert them into a more functional range of values.
   c. Choses members (i.e. the parents) based on their fitness value.
   d. Moves certain individuals directly to the next generation; these are called élite and are formed by some of the individuals with the lower fitness value in the current population.
   e. Creates children from the parents. Children are made with two methods: by *mutation* (it inserts random changes into a single parent's genes) or by *crossover* (it combines two parents' genes).
   f. Substitutes the current population with the children to form the next generation.

3. The algorithm stops when one of the stopping criteria is met
   (The MathWorks, Inc. (2014))

As we have just seen, there are three ways to create the next population; among them, the most interesting are the crossover and the mutation techniques:

- *Crossover children* are "the result of the combination of pairs of parents in the current population. At each coordinate of the child vector, the default crossover function randomly selects an entry, or *gene*, at the same coordinate from one of the two parents and assigns it

to the child. For problems with linear constraints, the default crossover function creates the child as a random weighted average of the parents" (The MathWorks, Inc. (2014)).

- *Mutation children* are "the result of a randomly change of the genes of individual parents. By default, for unconstrained problems the algorithm adds a random vector from a Gaussian distribution to the parent. For bounded or linearly constrained problems, the child remains feasible" (The MathWorks, Inc. (2014)).

Both processes are essential to the genetic algorithm: *Crossover* permits the algorithm to remove the best genes from different parents and recombine them into possibly superior children, while, *Mutation* increases the diversity of a population, enlarging the probability of the algorithm to generate individuals with a better fitness value. (The MathWorks, Inc. (2014))

As we said before, one of the most important elements for the genetic algorithm is the *diversity* of the population; in fact, it determines its performance; consequently, getting the right amount of diversity is extremely important: it cannot be too high or too low.

Another parameter that enters the concept of *diversity* is the *population size*, which determines the size of the population at each generation. Increasing the population size permits the genetic algorithm to search more points and thus find a better result. However, we should find the right population size: not too small and not too large; in fact the larger the population size, the longer the genetic algorithm takes to compute each generation. (The MathWorks, Inc. (2014))

The reason why we are interested on the genetic algorithm is that it can help us to find out the global minimum, instead of a simple local minimum (result that sometimes is not easy to find by optimization algorithms). The genetic algorithm can (sometimes) help us to reach this purpose, but only if we implement the right settings and strategies.

Among them, we can find an interesting strategy (i.e. the one I used in previous Section) called hybrid function. It is "an optimization function that runs after the genetic algorithm terminates, in order to improve the value of the fitness function. The hybrid function uses the final point from the genetic algorithm as its initial point" (The MathWorks, Inc. (2014)).

Obviously, inserting the right setting, when running the genetic algorithm, is important too: setting the right mutation coefficient, the rate at which the average amount of mutation

decreases, the diversity, the population size and the stopping criteria can help to find out the global minimum we are interested.

FOURTH CHAPTER: MY MODEL

4.1 Introduction

The purpose of my thesis is to reproduce/test one CAViaR specification and to compare the results obtained from the estimation of new models made from that CAViaR specification for three comparable portfolios.

In this chapter, I will introduce these portfolios and describes their composition.

I will then illustrate the models I constructed and present some related computational issues, trying, at the end, to give few possible solutions.

4.2 The portfolio

To choose my portfolios, I decided to focus my attention on some specific investment funds; for this reason, I decided to use the website Morningstar.it to select them.

I tried to choose three funds that could be comparable to each other; for doing so, I filtered all the possible funds by taking a specific category and benchmark.

At the end of my research, I decided to focus my attention on the "Stock European Large Cap Blend" category and on the "MSCI Europe NR EUR" index; consequently, I picked these three European investment funds:

1) "Capital Group European Growth and Income Fund (LUX)" (Class A4 EUR, ISIN LU0342012266, SEDOL B2NV7D5), launched October 30, 2002.
2) "Henderson Horizon Fund - Pan European Equity Fund" (Class A2 EUR, ISIN LU0138821268), launched November 30, 2001.
3) "JPM Europe Dynamic" (Class A (dist) EUR, ISIN LU0119062650), launched December 08, 2000.

All the information I collected on the funds come directly from Morningstar.it and the KIID (Key Investor Information Document) of Capital Group, Henderson Horizon and JPMorgan (i.e. the firms who launched the funds); most of the information are accurate as of 31 December 2015.

The common purpose of the three funds is to gain the capital growth in the long term by investing on companies established and/or that work mainly in Europe. Accordingly, the funds are more suited to investors with a long-term investment perspective (http://www.morningstar.it). In addition, all the funds can use derivatives for hedging or for a more efficient management.

Despite the common features, the funds differ for certain characteristics:

- The first one does not use tools as short selling and leverage; in addition, the management of the fund is based on the evaluation of every single stock, without any consideration on the weighting of any index;
- The second one is more flexible with respect to the possibility to concentrate its investments inside the European area; indeed, the fund invests at least 75% of its total capital in stocks of firms established in the European Economic Area (EEA).
- The third one applies an investment process based on the purchase of stocks that respect some specific style characteristics, as value, quality, price evolutions and profit trends. Indeed, the managers tend to gain the capital growth by exploiting psychological factors inherent in the stock markets.

The information about the returns of the investment funds are graphically summarize by these figures:

Figure 2. Returns of Capital Group European Growth and Income Fund (red line)

http://www.morningstar.it

Figure 3. Returns of Henderson Horizon Fund - Pan European Equity Fund (red line)

http://www.morningstar.it

Figure 4. Returns of JPM Europe Dynamic Fund (red line)

http://www.morningstar.it

These graphs contain three lines: the red one is the fund, the green one the benchmark of the category (i.e. MSCI Europe NR EUR) and the orange line is the category (i.e. stock European Large Cap Blend funds).

For having a better comparison among the category and the benchmark, I will also give a numerical comparison by using the following tables:

Table 2. First fund annual returns % (EUR)

Year	2008	2009	2010	2011	2012	2013	2014	2015
Total return	-43,14	37,28	7,68	-6,94	20,03	28,89	14,35	10,65
+/- Category	0,15	8,17	-3,95	4,06	2,22	9,28	9,07	-0,15
+/- Index	0,51	5,69	-3,42	1,14	2,73	9,07	7,51	2,42
Percentile in category	41	11	78	16	28	5	2	46

Table 3. Second fund annual returns % (EUR)

Year	2008	2009	2010	2011	2012	2013	2014	2015
Total return	-32,54	24,81	13,27	-7,65	19,4	25,17	6,38	15,3
+/- Category	10,75	-4,3	1,65	3,35	1,59	5,55	1,09	4,51
+/- Index	11,11	-6,79	2,18	0,43	2,1	5,34	-0,46	7,08
Percentile in category	7	79	34	20	33	10	36	14

Table 4. Third fund annual returns % (EUR)

Year	2008	2009	2010	2011	2012	2013	2014	2015
Total return	-46,66	29,12	15,38	-11,66	23,02	32,72	5,53	12,46
+/- Category	-3,37	0,01	3,76	-0,66	5,21	13,11	0,25	1,67
+/- Index	-3,01	-2,47	4,29	-3,58	5,73	12,9	-1,31	4,23
Percentile in category	79	51	23	60	9	3	48	33

As we can see from the table, during the last nine years the funds over-performed both the benchmark and the category for several years (with the exception of the third fund that over-performed the benchmark only half of the times).

Obviously, the high returns are associated with high risks, as we can see from the table below:

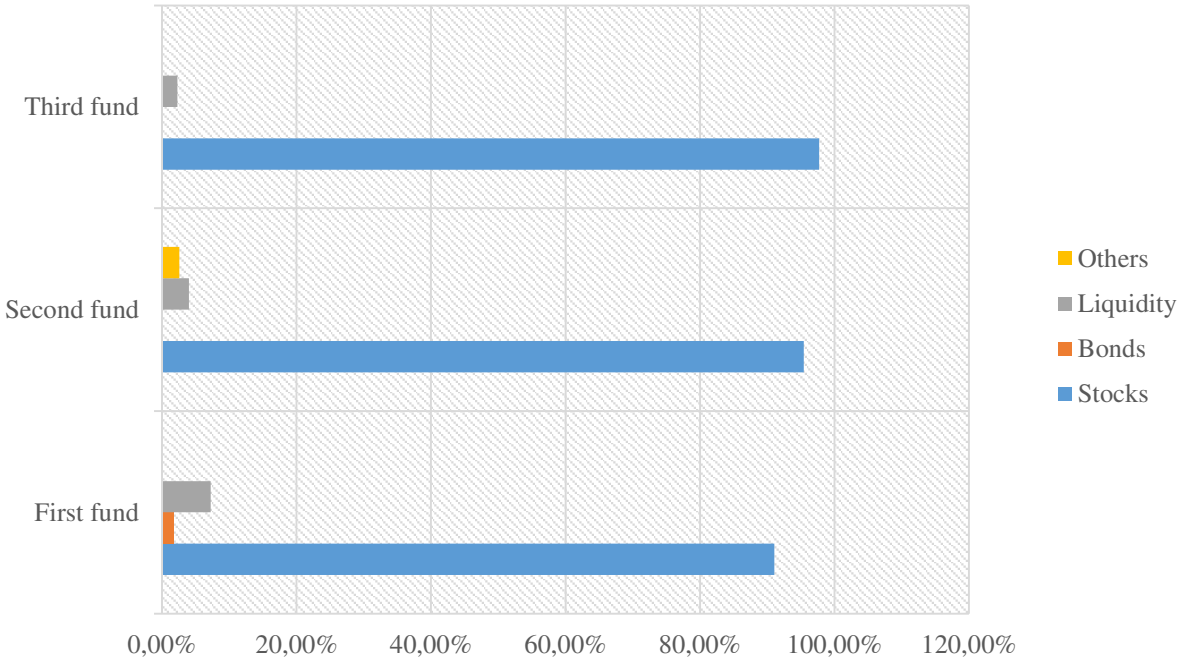
Table 5. Volatility Measures of the funds (3 years)

	First fund	Second fund	Third fund
Standard Deviation	12,00%	11,95%	12,47%
Mean return	18,52%	16,16%	17,21%
Sharpe ratio	1,42	1,26	1,28

All the three funds present a huge performance, associated with a huge standard deviation; this is mainly due to their composition: indeed, the funds invest most of their capital in stocks.

This is well shown by the following figure:

Figure 5. Investment preferences of the funds



Let now understand where are concentrated all the stock investments (i.e. in which areas and sectors); for this purpose, we can use some figures to visualize them:

Figure 6. Investment areas of the funds

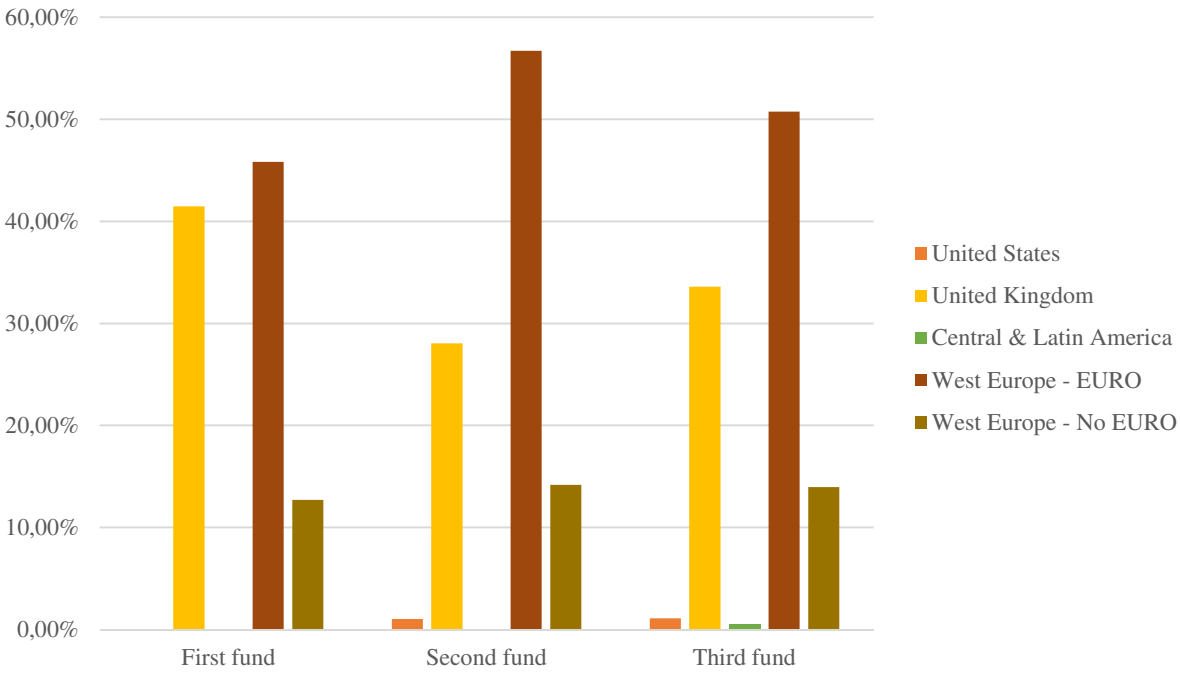
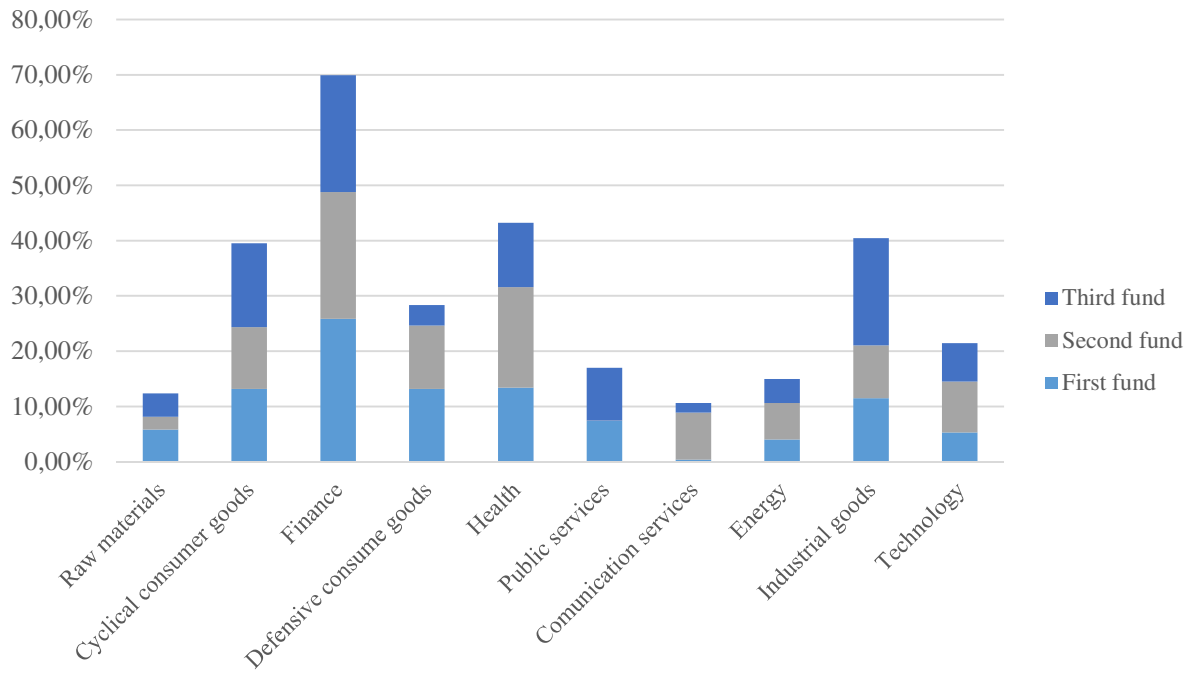


Figure 7. Investment sectors of the funds



4.3 My model

For creating the final model, it is essential to find out the risk factors connected to our investment funds.

The very first variable to use as risk factor is, obviously, the benchmark: for this reason, I decided to take the "MSCI EMU" that differs very little from Morningstar's benchmark, but has a longer data set.

For finding other risk factors, it is important to remember the core strategy of the portfolios we focused on: in fact, the funds invests mainly in stocks of large companies that work inside Europe. Consequently, some important risk factors could be the volatility inside the Europe, the price of the commodities and the EUR/US exchange rate (for imports/exports).

Moreover, the final graph of previous Section can help us to add new variables: indeed, the fact that all the three funds invests most of their capital in the financial sector (respectively 25,87%, 22,91% and 21,14%) suggests us that variables as interest rates (e.g. the Euribor, the Euro rate, etc.) can be suitable risk factors.

The second step for building our model is to understand what we want to analyze; in fact, CAViaR models can be used to investigate the data from different perspective. For this reason, I chose to focus on forecasting the future results: by manipulating the models, I can investigate

the data at time t, and give a value of the (future) VaR at time t+1. Computationally, this manipulation consists in lagging all the new variables of my model.

Another purpose that I want to reach in my thesis is to understand if, starting from one Engle and Manganelli's CAViaR specification, I can improve it by adding new variables (i.e. the risk factors).

For this reason, I will create and analyze different CAViaR models/specifications:

1) The first one, in which I will use as variable just the portfolio returns and I will estimate the original *SAV* specification;
2) The second one will be made by adding to the "original" model an important risk factor (i.e. the benchmark);
3) The last one will be composed by adding to the second model all the other risk factors I chose.

The last two models are just an "extension" of the original CAViaR model that I have computed in the previous Chapter; thus, all the computational aspects are the same as above. Nevertheless, some computational issue can still arise, especially in the third model where the number of variables increase a lot.

Therefore, let now focus on the new models and the issues introduced by the variables.

The first model is easy and it is essentially equal to the model we estimated in the previous Chapter. In particular, for all the three portfolios, I took a sample of 3,418 daily prices from Bloomberg; the samples covered a period of approximately 15 years (from October 30, 2002, to December 7, 2015).

All the computational aspect are the same we have already seen, with the exception of the way I initialize the optimization algorithm: as discussed above, we can adopt different implementation to make the algorithm running faster, but only two of them are the most efficient and they just differ on the technique they use to initialize the algorithm.

These approaches are the one suggested by Manganelli in one of his paper ((“VAR for VaR: Measuring tail dependence using multivariate regression quantiles”, 2015) and the one that use the Genetic Algorithm; consequently, from now on, I will use just them.

The results for the three funds are shown in the tables below:

Table 6. First Model 1% & 5% CAViaR estimations – First Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.0749	0.0345	-0.0264	0.0129
AutoregQuant	0.8948	2.4117e-178	0.9187	7.1475e-254
abs(y)	-0.2469	5.3142e-04	-0.1408	0.0017

Table 7. First Model 1% & 5% CAViaR estimations – Second Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.1608	0.0479	-0.0559	0.0117
AutoregQuant	0.8623	4.8203e-39	0.9013	0
abs(y)	-0.2874	0.0151	-0.1395	4.6180e-14

Table 8. First Model 1% & 5% CAViaR estimations – Third Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.1339	0.0013	-0.0572	0.0022
AutoregQuant	0.8789	1.1401e-194	0.8913	5.7883e-262
abs(y)	-0.2449	8.7870e-06	-0.1707	7.3360e-05

In this case (i.e. in the computation of the original CAViaR), I saw that the genetic algorithm is more efficient, because it gives me the same results in less time. For this reason, I will use it to make all the computations concerning the first model.

The second model adds the lagged benchmark variable to the first model. In this case, we proceed the same way as before: we compute the model by using the two different way to initialize the algorithm, but, from the results, we understand that the genetic algorithm cannot reach the global minimum. In fact, it tends not to give similar results every time I run it; on the contrary, the results seem to be more stable every time I use Manganelli's model. Thus, the last method appears to come closer to the global minimum.

Table 9. Second Model 1% & 5% CAViaR estimations – Manganelli's approach – First Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.1246	9.8885e-04	-0.0399	7.9537e-06
AutoregQuant	0.8651	0	0.9235	0
abs(y)	-0.2683	1.8375e-07	-0.1115	4.3996e-04
benchmark	0.2734	5.8690e-09	0.1171	9.7359e-17

Table 10. Second Model 1% & 5% CAViaR estimations – Manganelli's approach – Second Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.2961	3.4264e-08	-0.3122	1.4014e-14
AutoregQuant	0.7411	2.4214e-52	0.6200	4.9033e-31
abs(y)	-0.3837	1.0479e-05	-0.3701	7.3829e-06
benchmark	0.3800	9.4846e-21	0.4271	1.8168e-21

Table 11. Second Model 1% & 5% CAViaR estimations – Manganelli's approach – Third Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.2443	6.0567e-05	-0.0843	3.5192e-06
AutoregQuant	0.7953	3.8262e-43	0.8852	0
abs(y)	-0.3959	0.0026	-0.1490	1.1957e-04
benchmark	0.3180	1.0049e-26	0.1715	1.5515e-11

This is certainly due to the nature of the genetic algorithm that is a heuristic method; consequently, it is not sure that it can give every time the best results. Nevertheless, as I said in the previous Chapter, this algorithm can sometimes reach the optimum if we choose the right settings and strategies.

For this reason, I created a new setting for the algorithm: instead of using just one initial vector, I asked the genetic algorithm to create five initial vectors. Then I computed, for every initial vector the regression quantile (RQ) function and I chose three vectors, which produced the lowest RQ criterion, as initial values for the optimization routine.

This new setting guarantees us to have as good results as Manganelli's are; this is shown in the following tables:

Table 12. Second Model 1% & 5% CAViaR estimations – G.A. approach – First Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.1245	0.0010	-0.0398	8.0554e-06
AutoregQuant	0.8651	0	0.9234	0
abs(y)	-0.2686	1.8448e-07	-0.1116	4.4057e-04
benchmark	0.2733	5.9609e-09	0.1171	9.8949e-17

Table 13. Second Model 1% & 5% CAViaR estimations – G.A. approach – Second Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.3200	2.4225e-09	-0.3121	1.9660e-14
AutoregQuant	0.7285	5.6955e-53	0.6169	2.6328e-30
abs(y)	-0.3903	5.7256e-06	-0.3749	6.0302e-06
benchmark	0.3953	6.1086e-23	0.4284	1.5328e-21

Table 14. Second Model 1% & 5% CAViaR estimations – G.A. approach – Third Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.2443	6.0579e-05	-0.0845	4.9845e-06
AutoregQuant	0.7953	3.8424e-43	0.8854	0
abs(y)	-0.3959	0.0026	-0.1487	1.2385e-04
benchmark	0.3180	1.0057e-26	0.1715	1.5156e-11

An interesting fact is that both the methods requires the same amount of time to complete the loops; for this reason, both the methodologies are well suited to make the computations for this model (I preferred to choose the genetic algorithm methodology).

The most interesting case is the third model: because of the huge amount of variables involved in this model, the computational aspect is more complicated,

In fact, for taking into account all the possible relevant risk factors, I chose to create the last model by adding forty-seven variables to the second model (all the variables were downloaded from Datastream). The consequence of this choice is that both the Genetic and Manganelli's algorithms fail to achieve the vector of parameters that correspond to the global minimum (I recognized it thanks to the p-values made from the t-tests and the fact that the coefficients of the parameters change every time I start the loops).

One solution to this issue could be to decrease the number of variables; but one can ask "Which variables we should maintain and which ones we should drop off?"

The answer to this question is easy: the most significant ones. Obviously, this answer rises a following natural issue: how can we decide which variables are more significant than the rest?

Fortunately, it exists a statistical method that can help us solving this problem: it is called Principal Components Analysis.

"Principal component analysis is a variable reduction procedure. It is useful when you have obtained data on a number of variables (possibly a large number of variables), and believe that there is some redundancy in those variables. In this case, redundancy means that some of the variables are correlated with one another, possibly because they are measuring the same construct. Because of this redundancy, you believe that it should be possible to reduce the observed variables into a smaller number of principal components (artificial variables) that will account for most of the variance in the observed variables" (SAS Institute Inc.).

In particular, this procedure is "useful when you need a data reduction procedure that makes no assumptions concerning an underlying causal structure that is responsible for covariation in the data…[in fact] principal components depend solely on the covariance matrix (or the correlation matrix) of the original independent variables. Their development does not require a multivariate normal assumption." (SAS Institute Inc.).

Therefore, even if we have many variables, which together reproduce the total variability of the model, by using this method we can find few principal components that can account for much of this volatility. If so, there is (almost) as much information in few principal components as there is in the original variables; consequently, we can replace the initial variables with the principal components.

In addition, this type of analysis can help us to find out some relationships that were previously hidden, thus, allowing for interpreting the data in a different way (Johnson, Wichern (1998)).

From a mathematical point of view, principal components are "particular linear combinations of the p random variables $X_1, X_2, \dots, X_p$. Geometrically, these linear combinations represent the selection of a new coordinate system obtained by rotating the original system with $X_1, X_2, \dots, X_p$

as the coordinate axes. The new axes represent the directions with maximum variability and provide a simpler and more parsimonious description of the covariance structure" (Johnson, Wichern (1998)).

4.4 Principal Components Analysis & the Third Model

My analysis on principal components focused on forty-seven variables and did not take into account the benchmark; in fact, as we saw from the results of second model, this variable is every time very significant both from a statistic and from an economic point of view.

I computed the principal components by using a Matlab function; the results suggested me to use five principal components as new variables for the third model, in fact, they can together explain approximately 90% of the total population variance.

To interpret each principal component is necessary to calculate the correlation matrix between the components and the original forty-seven variables; in addition, we also have to compute the significance of every elements inside the correlation matrix (results are shown in Appendix B) and the total variance every principal components can explain.

Before analyzing the principal components is important to set some rules: these are necessary to separate the significant elements from the non-significant ones.
For this purpose, we can rely both on the p-value matrix (deciding to take the elements with a p-value<0.01) or on a constraint set on the correlation matrix (deciding to take just the elements that have a value greater or equal to 50%).
To better interpret the principal components, I chose to use the 50% constraint; in fact, the number of elements that have a significant p-value is too large to help us making any discrimination.

Let now analyze the results:

1) The first principal component explains 39.73% of the total variance in the "original" independent variables. The only variables that satisfy the 50% constraints are the S&P GSCI Commodity index, the crude oil prices, the VSTOXX and a similar volatility index for the US market (which are marked in the tables using the yellow color)

The first two are positively correlated with the principal component while the second two have a negative correlation; thus, the first principal component can be seen as a measure of stability in the market: in fact, it rises when the commodities grow and decreases when the volatilities grow. In particular, when the prices of all the commodities in the market grow, this indicates a situation of stability and growth in the economy, while an increase in the volatility indicates a situation of instability in the markets.

2) The second principal component explains 23.81% of the total variance of the independent variables. The significant variables that constitute the second principal components are the crude oil prices, the VSTOXX and a similar volatility index for the US market (they are marked in yellow).

   All these variables are positively correlated with the principal component; thus, this component could be interpret as a measure of instability. In fact, it rises when both the price of oil and the volatility grow (typical situation of instability in the markets).

3) The third principal component explains about 10.6% of the total volatility. By using the 50% constraint on the correlation elements, we end up with several significant variables that are mainly composed by many interest rate price indices and few interest rates spreads (yellow color).

   Consequently, this principal component can be interpreted as a measure of interest rates levels in the period we took (from October 30, 2002, to December 7, 2015).

4) When we filter the forth principal component, we do not find any significant variables. In fact, the most weighted variables reach just a weight of 45.67% (the VSTOXX) and of -41.67% (volatility index for the US market); these are marked in the table using a brown color. This is not a big issue, in fact, this principal component explains only 8.27% of the total variability of the independent variables, and thus it is not very relevant for the model.

5) The last principal component contains only 6.55% of the total volatility and it is mainly composed by the EUR/USD exchange rate and the S&P GSCI Commodity index (yellow color), which are both positively correlated with the principal component. We could interpret it as a price of the inputs and, thus, a measure of competition of the European economy; in fact, the price of euro and inputs are two of the most important variables for the profitability of European big firms.

Thanks to the principal components technique we reduced the number of independent variables to be added in our third model; in fact, now, the model must find just a total amount of parameters equal to nine, thus, now there are more chances that our algorithms can come closer the global minimum point.

As, we discussed above, the genetic algorithm can be a good method for finding out the global minimum only if we set it correctly; for this reason, I created another new setting for the genetic algorithm: I increased the number of initial vectors (from five to fifteen). Then, as I did above, I computed for every initial vector the regression quantile (RQ) function and I chose five vectors (instead of three), which produced the lowest RQ criterion, as initial values for the optimization routine.
This new setting improved the results, but, as I said, we are never sure that a heuristic method can reach the global minimum; for this reason, I implemented, for every fund of investment, some comparisons between the new Genetic Algorithm and Manganelli's algorithm:

1) For the first fund, I saw that the two methods give me approximately the same results, with the difference that the genetic algorithm spent less time than Manganelli's one (around five seconds less); thus, I chose to use the genetic approach to compute the third model (and its future estimations) for the first fund. The results are shown in the table below:

Table 15. Third Model 1% & 5% CAViaR estimations – G.A. approach – First Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.0900	0.0035	-0.0317	2.0122e-04
AutoregQuant	0.8956	1.4867e-121	0.9262	0
abs(y)	-0.1957	0.0074	-0.1102	2.9158e-05
benchmark	0.0883	0.0776	0.0650	0.0036
component 1	0.0468	0.0168	0.0403	2.3834e-04
component 2	-0.0718	0.0010	-0.0639	1.8125e-06
component 3	0.0122	0.3572	0.0474	6.5900e-05
component 4	0.1099	0.0785	0.1477	1.3193e-05
component 5	-0.0687	0.1158	-0.0302	0.0497

2) For the second fund, I noticed that the genetic algorithm gives me "weak" results (in fact, by running the algorithm several times, I saw that they are not always close to

Manganelli's outcomes). Consequently, I preferred to use Manganelli's method, even if it requires more time to be computed (about 8 seconds more). We can see the results in the following table:

Table 16. Third Model 1% & 5% CAViaR estimations – Manganelli's approach – Second Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.1493	1.4301e-06	-0.4232	3.9580e-11
AutoregQuant	0.8252	4.2530e-177	0.4586	6.9943e-19
abs(y)	-0.3271	1.5284e-04	-0.4681	3.8566e-30
benchmark	0.1580	0.0053	0.4725	1.4864e-11
component 1	0.0893	3.4521e-08	0.0124	0.3382
component 2	-0.1135	5.2343e-07	-0.0272	0.1892
component 3	0.0499	0.0145	0.0541	0.0199
component 4	0.1998	9.6402e-04	0.2742	3.1042e-13
component 5	-0.1320	0.0021	-0.1396	8.2479e-05

3) Similar conclusions can be made for the last fund: even if the genetic algorithm takes less time (seven seconds less), it is not as precise as Manganelli's loops. Consequently, I chose to use it for making all the computations for the third model on the third fund. The table below shows the results:

Table 17. Third Model 1% & 5% CAViaR estimations – Manganelli's approach – Third Fund

	BetaHat1	CoeffPvalue1	BetaHat5	CoeffPvalue5
Costant	-0.3402	0.0013	-0.0867	1.9149e-05
AutoregQuant	0.7435	7.2597e-45	0.8546	5.1289e-149
abs(y)	-0.3761	1.6863e-05	-0.2093	2.7773e-04
benchmark	0.1889	0.0296	0.1490	9.5409e-08
component 1	0.0588	4.7472e-04	0.0279	0.0088
component 2	-0.1125	0.0033	-0.0576	9.3650e-06
component 3	0.0870	0.0169	0.0637	0.0071
component 4	0.3744	2.3361e-04	0.2338	4.0374e-06
component 5	-0.1766	6.5880e-04	-0.0978	3.9018e-05

Before concluding this chapter, I would like to highlights few things I consider relevant:

1) It is not easy to find the global minimum, especially when we increase the number of independent variables; in this case, most of the algorithms fail or require a huge amount of time. Sometimes we must find a compromise that can give us a local minimum quite close to the global one and that does not require us an exaggerate amount of time.
2) The genetic algorithm is a heuristic method that can help us to reach the optimal solution very quickly, but it requires the right setting, which imply many experiments and comparisons. Nevertheless, as we saw, it can also fail and alternative methods can be preferred.

FIFTH CHAPTER: EMPIRICAL RESULTS

5.1 Introduction

In this Chapter, I will describe and analyze the final computations I made.
In the first part, I will introduce the technique that permitted me to investigate the data in a more rigorous fashion; then I will present a test about the accuracy of my models.
Lastly, I will introduce some issues and give some overall conclusions on the analysis.

5.2 The rolling window method

As we have seen in the previous chapter, the usage of an entire time-series data set gives us just one estimation of the parameters we are interested; this is not very useful if we want to make some comparisons and analysis among different models and different portfolios. In fact, the use of just one estimation of parameters for each model and each fund can lead to misinterpret the results; consequently, for making rigorous analysis we need to estimate a bigger amount of parameters.
For doing this, we can use a rolling window method.
This technique consists in using a time window and making the estimation of the parameters on this "restricted" time series data set; then, after the computation, we move the time window further in the data set and estimate another set of parameters. We repeat this procedure until the last data contained in the window is equal to the last data contained in the entire data set.

In my case, for each fund and model, I have a times series consisting of 3,418 observations for each variable and, consequently, by using a rolling window of 1000 observations, I end up with 2,418 vectors of estimated parameters. This huge amount of estimated parameters can help us to better analyze the results, but, as I mentioned several times in my thesis, we have also to pay attention to the amount of time we spend for reaching the outcomes.
For this reason, I decided not to use all the 2,418 rolling windows for computing the vectors of parameters, but only 484 windows: in fact, for every five windows, I estimated only the vector of parameters of the first window, extending it to the other following four windows. This permitted me to calculate the quantiles and the p-values (on the estimated parameters) for every window, thus, reducing remarkably the amount of time spent.

The negative side of this methodology is that it gives us less estimated parameters and that it is not completely precise: in fact, it extends the estimated coefficients of an initial time window to the following four windows for computing their quantiles and p-values.

Consequently, also in this case, we have an issue concerning the tradeoff between accuracy and timing of estimations. Fortunately, our methodology is jet a good compromise between precision and time saving: extending a coefficients estimation to four following windows does not distort dramatically the calculation of quantiles and p-values. Thus, we can use this methodology without problems with the huge advantage of saving a lot of time in the computations.

For a better analysis of the results, I created some tables in which I summarized the most interesting features of the outcomes; in particular, the first two columns of the tables concern the number of estimated coefficients and the number of significant estimated coefficients (i.e. those that have a p-value less than 1%). The other columns refer to the significant coefficients and include the mean, the standard deviation, the minimum value, the maximum value, the number of positive coefficients and the number of negative coefficients.
All the results are presented in Appendix C.

As we can see from the table (yellow color), the autoregressive quantile coefficients seem to be the most stable in all the models and for every fund: the number of significant coefficients on the number of estimated coefficients is every time very high (at least higher than 80%). On the contrary, the other variables are much more unstable: few of them reach a percentage of "significance" higher than 50% and, moreover, this result does not persist for every model and for every fund (this conclusion is shown in the table using the brown color).

These outcomes permit us to make some first conclusions: the fact that the coefficients of the autoregressive term are always very significant permits us to say "that the phenomenon of clustering of volatilities is relevant also in the tails" (Engle, Manganelli (2002)).

We now turn into a second analysis that allows us to understand if our model are precise or not; for doing so, we are going to use a test that was entirely created by Engle and Manganelli: the so-called Dynamic Quantile (DQ) test.

5.3 The DQ-Test

If our models are the true DGP (Data Generating Process), "then $\Pr\big(y_t < f_t(\beta^0)\big) = \theta$. This is equivalent to requiring that the sequence of indicator functions $\{I_t\}_{t=1}^{T} = \big\{I\big(y_t < f_t(\beta^0)\big)\big\}_{t=1}^{T}$ be independent and identically distributed. Hence, a property that any VaR estimate should satisfy is that of providing a filter to transform a (possibly) serially correlated and heteroskedastic time series into a serially independent sequence of indicator functions" (Engle, Manganelli (2002)).

For doing this, Engle and Manganelli (2002) proposed a test, which has better power properties than other existing tests: the Dynamic Quantile Test (DQ-Test). In fact, even if most of the tests can "detect the presence of serial correlation in the sequence of indicator functions $\{I_t\}_{t=1}^{T}$, this is only a necessary, but not sufficient condition to assess the performance of a quantile model. Indeed, it is not difficult to generate a sequence of independent $\{I_t\}_{t=1}^{T}$ from a given sequence of $\{y_t\}_{t=1}^{T}$" (Engle, Manganelli (2002)).

Indeed, as was pointed out by Allen and Singh (2010), "A relevant VaR model should also feature a sequence of VaR violations which are not serially correlated".

Engle and Manganelli achieved this result by using a *Hit* function of this type:

$$Hit_t(\beta^0) = I\big(y_t < f_t(\beta^0)\big) - \theta$$

Where the function $Hit_t(\beta^0)$ takes a value $(1-\theta)$ every time y_t falls below the quantile, and it takes the value $-\theta$ in all the other circumstances. The equation above indicates that the expectation of $Hit_t(\beta^0)$ is equal to zero. Additionally, based on the definition of the quantile "we also assume that the conditional expectation of $Hit_t(\beta^0)$ given a set of information at period $t-1$ is zero. This implies that $Hit_t(\beta^0)$ must be uncorrelated with its own lagged values as well as with $f_t(\beta^0)$ and its expected value should equal zero. If these assumptions hold for $Hit_t(\beta^0)$ then we are certain that we have no misspecification error introduced, there is no autocorrelation in the *hits*, and we will obtain the correct fraction of exceptions" (Kouretas, Zarangas (2005)).

Based on these considerations, Engle and Manganelli (2002) proposed the following test:

If T symbolizes the number of in sample observations and N the number of out of sample observations, a natural way to build a test is to check whether the test statistic $X_N'(\hat{\beta}_T)Hit_N(\hat{\beta}_T)$ is significantly different from zero, where $X_n'(\hat{\beta}_T),\ n = T+1, \dots, T+N$.

The typical row of $X_N(\hat{\beta})$ (possibly depending on $\hat{\beta}$), is a q-vector measurable-Ω_n and $Hit_N(\hat{\beta}) = [Hit_{T+1}(\hat{\beta}), \dots, Hit_{T+N}(\hat{\beta})]'$.

Accordingly, the two researchers proposed this out of the sample version of the Dynamic Quantile (DQ) test statistic:

$$DQ_0 \equiv \frac{Hit_N'(\hat{\beta}_T)X_N(\hat{\beta}_T)\left[X_N'(\hat{\beta}_T)X_N(\hat{\beta}_T)\right]^{-1}X_N'(\hat{\beta}_T)Hit_N(\hat{\beta}_T)}{\theta(1-\theta)} \underset{\substack{T\to\infty \\ N\to\infty}}{\overset{d}{\to}} \chi_q^2$$

provided that $X_N'(\hat{\beta}_T)X_N(\hat{\beta}_T)$ is nonsingular. The limit for $T \to \infty$ is required to ensure that $\hat{\beta}_T \overset{p}{\to} \beta^0$ (Engle and Manganelli (2002)).

This test is very useful, in fact as Engle and Manganelli (2002) revealed, it "can be used by regulators to check whether the VaR estimates submitted by a financial institution satisfy some basic requirements every good quantile estimates must have, such as unbiasedness, independent hits and independence of the quantile estimates. [Additionally,] the nicest features of the out-of-sample DQ test are its simplicity and the fact that it does not depend on the estimation procedure: to implement it, the evaluator (either the regulator or the risk manager) just needs a sequence of VaR's and the corresponding values of the portfolio".

Let now move to the computational aspect; from this point of view I completely followed Engle and Manganelli's method with two main exceptions:

1. The instruments used in the out-of-sample DQ test were the lagged returns, the square of the lagged returns, the lagged VaR forecast and the first four lagged hits (instead of a constant, the VaR forecast and the first four lagged hits, as used by Engle and Manganelli).
2. For creating the out-of-sample DQ test, Engle and Manganelli divided the historical data into two parts: they used the first 2,892 observations to estimate the model and the last 500 for out-of-sample testing. I proceeded in a similar but not the same fashion: by using the rolling method seen above, I estimated 2,418 vectors of parameters and, for each estimation, I created the future quantile. In other words, for the period t+m I can obtain

the quantile for t+m+1; this is possible because all my independent variables are lagged (thus I just need the information set at t+m for estimating the t+m+1 quantile). Then I aggregated the series of t+m+1 quantiles (obtained for different values of t) and on this sequence I computed the out-of-sample DQ-test.

The results for the percentage of times the VaR is exceeded and the p-value of the Dynamic Quantile tests are shown in the following tables:

Table 18. Hits and DQ-test for out-of-sample 1% VaR

	FirstFund	SecondFund	ThirdFund
HitOut(%)- First Model	1.0753	1.7370	1.1993
DQ-test OUT (p-value) - First Model	1.4871e-04	0	1.3784e-08
HitOut(%)- Second Model	1.5715	1.6543	1.5715
DQ-test OUT (p-value) - Second Model	8.4422e-11	4.0629e-10	1.3250e-07
HitOut(%)- Third Model	6.1208	9.0984	1.8610
DQ-test OUT (p-value) - Third Model	0	0	0

Table 19. Hits and DQ-test for out-of-sample 5% VaR

	FirstFund	SecondFund	ThirdFund
HitOut(%)- First Model	5.2936	5.8313	5.0455
DQ-test OUT (p-value) - First Model	1.3209e-09	2.8388e-12	0
HitOut(%)- Second Model	5.4177	5.6658	5.1282
DQ-test OUT (p-value) - Second Model	0	0	1.1102e-16
HitOut(%)- Third Model	11.4971	15.0124	6.2862
DQ-test OUT (p-value) - Third Model	0	0	0

As we can see from the results, all the models are very precise (according to the p-values of DQ-tests), even if some of them exceeds the 1% or 5% of hits (especially the first and second fund in the third model - both for the 1% and 5% VaR case). “This shows that looking only at the number of exceptions (as suggested by the Basle Committee on Banking Supervision

(1996)) may be a very unsatisfactory way of evaluating the performance of a VaR model" (Engle, Manganelli (2002)).

5.4 Crossings and Conclusions

Let now turn into a very relevant issue concerning the quantiles, i.e. the quantiles crossings. The problem is well explained by Bondell, Reich and Wang (2010): "when an investigator wishes to use quantile regression at multiple percentiles, the quantile curves can cross, leading to an invalid distribution for the response. Given a set of covariates, it may turn out, for example, that the predicted 95th percentile of the response is smaller than the 90th percentile, which is impossible".

This is a big issue and should be taken into consideration while analyzing our results, because it can lead to misinterpret them; for this reason, I computed the number of crossing between the 1% and 5% quantiles for each model and each fund. The results are shown in the following table:

Table 20. Crossing between quantiles

	FirstFund	SecondFund	ThirdFund	NumberOfSeries
1stModel	41	73	182	2418
2ndModel	1403	1800	1224	2418
3rdModel	2278	2367	2383	2418

As we can see from the table, the number of crossing is very low in the first model, while it "explodes" for the second and third models; these outcomes suggest us to discard the second and third model, in fact they produce biased results.
It is interesting to note that this last result is shared by all the three funds.

In view of the above, we could make some conclusions:

1) It seems that, for my data sets, only the CAViaR specification introduced by Engle and Manganelli can survive all the computations made above; thus, I would say that creating

new models by inserting risk factors inside the "original" model could just make it worse.

2) If we analyze the results in Appendix C just focusing on the first model, we can notice that only the autoregressive quantile component is very significant (both in the 1% and in 5% VaR). In fact, the constant term and the absolute value parameter are not such considerable (in particular, the absolute value estimations improve moving from 1% to 5% VaR, but they never become significant). This fact could suggest that our "original" CAViaR model is not well suited for our data.

These two results introduce an interesting consideration: there are different kind of processes for modelling the CAViaR (we saw just few of them in the Second Chapter) and it does not exist one that always outperforms the others: the best model depends on the data set and on the quantiles. This observation implies an obvious consideration: to reach the best model we have to compare different models any time we change the portfolio/s and the quantiles we are investigating.

These considerations are in line with Engle and Manganelli's outputs (2002): in fact, the two researchers performed all the models seen in Chapter 2, demonstrating that "the DQ tests select different models for different confidence levels". This result is very important and led to the conclusion "that the process governing the tail behavior might change as we move further out in the tail. In particular, this contradicts the assumption behind GARCH and RiskMetrics, since these approaches implicitly assume that the tails follow the same process as the rest of the returns. While GARCH might be a useful model to describe the evolution of volatility, the results in this paper show that it might provide an unsatisfactory approximation when applied to tail estimation" (Engle, Manganelli (2002)).

SUMMARY AND CONCLUDING REMARKS

In my thesis, I introduced a new structure for estimating the VaR of portfolios returns.

This procedure has been firstly introduced by Engle and Manganelli (1999) and it is a semiparametric method "which shifts the analysis of developing a good measure of the VaR from the distribution of the portfolio returns directly to the behavior of the quantile" (Kouretas, Zarangas (2005)). For doing so, the two researchers proposed the so-called Conditional Autoregressive Value at Risk (CAViaR) models, "which specify the evolution of the quantile over time using a special type of autoregressive process" (Engle, Manganelli (2002)), utilizing a minimization regression quantiles loss function to estimate the unknown parameters.

Starting from this basis, I chose to focus my attention on one CAViaR specification introduced by the two researchers (i.e. the *Symmetric Absolute Value*) studying its behavior applied to three sets of comparable portfolio returns. I also tried to improve the basic model by adding different risk factors, ending up with three CAViaR models for each fund, which I compared to find out the best VaR predictor model.

The overall results show that the original CAViaR specification cannot be improved by adding risk factors, representing, thus, the best model we found for our data sets. Another important outcome is that only one estimated coefficient of the original model is very significant, suggesting us that maybe the original CAViaR specification does not perfectly explain our data sets of portfolio returns.

These conclusions are very interesting and important for the financial sector: the fact that it does not exist only one right CAViaR specification should suggest the regulators and financial firms not to overlook the estimation of VaR, which, when it is accurate, requires a lot of experiments and comparisons to be found.

*Appendix A: Assumptions**

Consistency Assumptions

C0. (Ω, F, P) is a complete probability space and $\{\varepsilon_{t\theta}, x_t\}$, $t = 1,2, \ldots$, are random vectors on this space.

C1. The function $f_t(\beta)$: $\mathbb{R}^{kt} \times B \to \mathbb{R}$ is such that for each $\beta \in B$, a compact subset of $\mathbb{R}^p$, $f_t(\beta)$ is measurable with respect to the information set Ω_t and $f_t(\cdot)$ is continuous in B, $t = 1,2, \ldots$, for a given choice of explanatory variables $y_{t-1}, x_{t-1}, \ldots, y_1, x_1$.

C2. Conditional on all the past information Ω_t, the error term $\varepsilon_{t\theta}$ form a stationary process, with continuous conditional density $h_t(\varepsilon|\Omega_t)$, and continuous joint density $h_t^{\varepsilon,\Omega}(\varepsilon, y, x)$.

C3. There exists $h > 0$ such that for all t $h_t(0|\Omega_t) \geq h$.

C4. $|f_t(\beta)| < K(\Omega_t)$ for each $\beta \in B$ and for all t, where $K(\Omega_t)$ is some (possibly) stochastic function of variables that belong to the information set, such that $E(|K(\Omega_t)|^r) \leq K_0 < \infty$, $r > 1$ for some constant K_0.

C5.$E(|\varepsilon_{t\theta}|) < \infty$ for all t.

C6. $\{[\theta - I(y_t < f_t(\beta))][y_t - f_t(\beta)]\}$ obeys the law of large numbers.

C7. For every $\delta > 0$, there exists a $\tau > 0$ such that if $||\beta - \beta^0|| \geq \delta$,

$$\liminf_{T\to\infty} T^{-1} \sum P[|f_t(\beta) - f_t(\beta^0)| > \tau] > 0.$$

Asymptotic Normality Assumptions

AN1. $f_t(\beta)$ is differentiable in B and for all β and γ in a neighborhood v_0 of β^0, such that $||\beta - \gamma|| \leq d$ for d sufficiently small and for all t.

(i) $||\nabla f_t(\beta)|| \leq F(\Omega_t)$, where $F(\Omega_t)$ is some (possibly) stochastic function of variables that belong to the information set and $E(|F(\Omega_t)|^3) \leq F_0 < \infty$, for some constant F_0

(ii) $\nabla f_t(\beta)$ satisfies the Lipschitz condition $||\nabla f_t(\beta) - \nabla f_t(\gamma)|| \leq M||\beta - \gamma||$, where $M < \infty$

AN2. (i) $h_t(\varepsilon|\Omega_t) \leq N < \infty, \forall t$

(ii) $h_t(\varepsilon|\Omega_t)$ satisfies the Lipschitz condition
$|h_t(\lambda_1|\Omega_t) - h_t(\lambda_2|\Omega_t)| \leq L|\lambda_1 - \lambda_2|$, where $L < \infty, \forall t$.

AN3. There exists $\delta > 0$, such that for any $\beta \in B$ and T sufficiently large

$\det(T^{-1} \sum_{t=1}^{T} E\,[\nabla' f_t(\beta) \nabla f_t(\beta)]) > \delta$.

AN4. $\{[\theta - I(y_t < f_t(\beta))]\nabla' f_t(\beta)\}$ satisfies a central limit theorem

Variance Covariance Matrix Estimation Assumptions

VC1. $\hat{c}_T/c_T \xrightarrow{p} 1$, where the nonstochastic sequence c_T satisfies $c_T = o(1)$ and $c_T^{-1} = o(T^{1/2})$.

VC2. $E(|F(\Omega_t)|^4) \leq F_1 < \infty$. For all t and for some constant F_1, where $F(\Omega_t)$ has been defined in AN1 (i).

* The present appendix was entirely formulated by Engle and Manganelli (2002)

Appendix B: Principal Components

Table 1. Correlation matrix	Pricipal Components				
US $ TO EURO (WMR&DS) - EXCHANGE RATE	0.1543	-0.1137	0.0702	0.0092	0.5197
S&P GSCI Commodity Total Return - RETURN IND. (OFCL)	0.7411	0.3335	0.0826	0.1956	0.5281
Crude Oil-Brent Dated FOB EUR/BBL	0.7397	0.6269	0.0136	-0.0546	-0.2309
BARCLAYS EURO AGG GOVERNMENT (E)	-0.2287	0.0901	0.8299	-0.1988	0.0552
BARCLAYS EURO AGGREGATE (E)	-0.2462	0.1136	0.8348	-0.1850	0.0424
BARCLAYS PAN-EUR. AGG (E)	-0.2976	0.1426	0.8349	-0.1529	0.0393
BD BENCHMARK 10 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	-0.3940	0.2263	0.7790	-0.0596	-0.0444
BOFA ML EUR UNION GVT. (E) - TOT RETURN IND	-0.2432	0.1342	0.8187	-0.1811	-0.0336
BD BENCHMARK 30 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	-0.3656	0.1950	0.8281	-0.1042	-0.0553
EMU BENCHMARK 30 YR. DS GOVT. INDEX - CLEAN PRICE INDEX	-0.3620	0.1927	0.8253	-0.1049	-0.0575
EMU BENCHMARK 10 YR. DS GOVT. INDEX - CLEAN PRICE INDEX	-0.3940	0.2262	0.7785	-0.0594	-0.0445
FR BENCHMARK 10 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	-0.2982	0.1436	0.8011	-0.1329	0.0216
IBOXX EURO CORP. ALL MATS - PRICE INDEX	-0.2029	0.1453	0.6789	-0.1439	0.0066
IBOXX EURO OVERALL INDEX ALL MATS. - PRICE INDEX	-0.2364	0.1070	0.8263	-0.1734	0.0349
IBOXX EURO SOVEREIGN EZONE ALL MATS - PRICE INDEX	-0.2183	0.0746	0.8442	-0.2033	0.0601
IT BENCHMARK 10 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	-0.0043	-0.0728	0.5477	-0.2496	0.1474
IT BENCHMARK 30 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	-0.0137	-0.0551	0.6424	-0.2895	0.1225
EURO SPOT WEEK FX VOL - EXCHANGE RATE	-0.2427	0.2587	-0.0489	-0.0261	-0.0870
Vstoxx	-0.7092	0.5269	0.0167	0.4567	-0.0327
volatility usa	-0.6288	0.5942	-0.2086	-0.4167	0.1774
EURO RATE 3 MONTH (DS SYNTHETIC) - OFFERED RATE	0.0109	0.0656	-0.0341	-0.0391	0.0385
EURO RATE 6 MONTH (DS SYNTHETIC) - OFFERED RATE	0.0541	0.0465	-0.0770	-0.0264	0.0259
EURO RATE 1 MONTH (DS SYNTHETIC) - OFFERED RATE	0.0161	0.0524	-0.0119	-0.0243	0.0256
EURO RATE 1 YEAR (DS SYNTHETIC) - OFFERED RATE	0.0725	0.0213	-0.1300	-0.0213	0.0102
(TR) EURO VS EURIBOR 6M IR SWAP 10Y - MIDDLE RATE	0.3362	-0.2010	-0.7081	0.1341	0.0291
EBF EURIBOR 3M DELAYED - OFFERED RATE	0.0087	0.0607	-0.0282	-0.0348	0.0415
EBF EURIBOR 6M DELAYED - OFFERED RATE	0.0571	0.0465	-0.0691	-0.0266	0.0203
IBA EUR IRS ISDAFIX 10Y DELAYED - MIDDLE RATE	0.1916	-0.1248	-0.3706	-0.0180	0.0315
IBA EUR IRS ISDAFIX 5Y DELAYED - MIDDLE RATE	0.1795	-0.1143	-0.3185	-0.0244	0.0174
INEU05(CM01)	0.1552	-0.0846	-0.2065	0.0004	-0.0012
INEU05(CM10)	0.2325	-0.1216	-0.5678	0.0950	0.0422
INEU05(CM02)	0.2360	-0.1504	-0.4041	0.0530	0.0102
INEU05(CM03)	0.2601	-0.1571	-0.5123	0.0603	0.0242
INEU05(CM05)	0.2849	-0.1629	-0.6340	0.0765	0.0205
EURO SPOT WEEK OIS - MIDDLE RATE	0.0689	-0.0172	-0.0203	-0.0135	0.0341
EURO 6 MONTH OIS - MIDDLE RATE	0.2823	-0.1907	-0.1910	0.0333	0.0560
EURO 1 YEAR OIS - MIDDLE RATE	0.2814	-0.1955	-0.3199	0.0664	0.0285
ECU EURO-ECU 5 MTH - MIDDLE RATE	-0.0135	0.0191	-0.0491	-0.0039	0.0179
ECU EURO-ECU 1 YEAR - MIDDLE RATE	0.0743	-0.0334	-0.1034	-0.0146	0.0219
IBA EUR IBK. LIBOR 1M DELAYED - OFFERED RATE	0.0156	0.0526	-0.0122	-0.0245	0.0252
IBA EUR IBK. LIBOR 6M DELAYED - OFFERED RATE	0.0534	0.0464	-0.0775	-0.0258	0.0260
IBA EUR IBK. LIBOR O/N DELAYED - OFFERED RATE	0.0295	0.0118	-0.0036	0.0193	0.0300
(TR) EURO VS EURIBOR 6M IR SWAP 5Y - MIDDLE RATE	0.3070	-0.1996	-0.6163	0.1275	0.0193
INEU03(CM01)	0.1944	-0.1059	-0.2453	0.0148	0.0022
INEU03(CM10)	0.2448	-0.1411	-0.6003	0.1003	0.0342
INEU03(CM05)	0.2845	-0.1627	-0.6269	0.0770	0.0222
INEU03(CM02)	0.2529	-0.1531	-0.4331	0.0463	0.0154

Table 2. Associated p-values	Pricipal Components				
US $ TO EURO (WMR&DS) - EXCHANGE RATE	0.0000	0.0000	0.0000	0.5919	0.0000
S&P GSCI Commodity Total Return - RETURN IND. (OFCL)	0	0.0000	0.0000	0.0000	0.0000
Crude Oil-Brent Dated FOB EUR/BBL	0	0	0.4281	0.0014	0.0000
BARCLAYS EURO AGG GOVERNMENT (E)	0.0000	0.0000	0	0.0000	0.0012
BARCLAYS EURO AGGREGATE (E)	0.0000	0.0000	0	0.0000	0.0131
BARCLAYS PAN-EUR. AGG (E)	0.0000	0.0000	0	0.0000	0.0215
BD BENCHMARK 10 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	0.0000	0.0000	0	0.0005	0.0094
BOFA ML EUR UNION GVT. (E) - TOT RETURN IND	0.0000	0.0000	0	0.0000	0.0496
BD BENCHMARK 30 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	0.0000	0.0000	0	0.0000	0.0012
EMU BENCHMARK 30 YR. DS GOVT. INDEX - CLEAN PRICE INDEX	0.0000	0.0000	0	0.0000	0.0008
EMU BENCHMARK 10 YR. DS GOVT. INDEX - CLEAN PRICE INDEX	0.0000	0.0000	0	0.0005	0.0093
FR BENCHMARK 10 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	0.0000	0.0000	0	0.0000	0.2066
IBOXX EURO CORP. ALL MATS - PRICE INDEX	0.0000	0.0000	0	0.0000	0.6995
IBOXX EURO OVERALL INDEX ALL MATS. - PRICE INDEX	0.0000	0.0000	0	0.0000	0.0416
IBOXX EURO SOVEREIGN EZONE ALL MATS - PRICE INDEX	0.0000	0.0000	0	0.0000	0.0004
IT BENCHMARK 10 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	0.8013	0.0000	0.0000	0.0000	0.0000
IT BENCHMARK 30 YEAR DS GOVT. INDEX - CLEAN PRICE INDEX	0.4217	0.0013	0	0.0000	0.0000
EURO SPOT WEEK FX VOL - EXCHANGE RATE	0.0000	0.0000	0.0043	0.1278	0.0000
Vstoxx	0	0.0000	0.3278	0.0000	0.0558
volatility usa	0	0	0.0000	0.0000	0.0000
EURO RATE 3 MONTH (DS SYNTHETIC) - OFFERED RATE	0.5222	0.0001	0.0464	0.0222	0.0243
EURO RATE 6 MONTH (DS SYNTHETIC) - OFFERED RATE	0.0016	0.0065	0.0000	0.1224	0.1302
EURO RATE 1 MONTH (DS SYNTHETIC) - OFFERED RATE	0.3470	0.0022	0.4885	0.1548	0.1351
EURO RATE 1 YEAR (DS SYNTHETIC) - OFFERED RATE	0.0000	0.2141	0.0000	0.2139	0.5506
(TR) EURO VS EURIBOR 6M IR SWAP 10Y - MIDDLE RATE	0.0000	0.0000	0	0.0000	0.0887
EBF EURIBOR 3M DELAYED - OFFERED RATE	0.6095	0.0004	0.0987	0.0417	0.0152
EBF EURIBOR 6M DELAYED - OFFERED RATE	0.0008	0.0066	0.0001	0.1194	0.2360
IBA EUR IRS ISDAFIX 10Y DELAYED - MIDDLE RATE	0.0000	0.0000	0.0000	0.2915	0.0655
IBA EUR IRS ISDAFIX 5Y DELAYED - MIDDLE RATE	0.0000	0.0000	0.0000	0.1546	0.3103
INEU05(CM01)	0.0000	0.0000	0.0000	0.9833	0.9461
INEU05(CM10)	0.0000	0.0000	0.0000	0.0000	0.0135
INEU05(CM02)	0.0000	0.0000	0.0000	0.0019	0.5514
INEU05(CM03)	0.0000	0.0000	0.0000	0.0004	0.1570
INEU05(CM05)	0.0000	0.0000	0	0.0000	0.2297
EURO SPOT WEEK OIS - MIDDLE RATE	0.0001	0.3150	0.2364	0.4287	0.0464
EURO 6 MONTH OIS - MIDDLE RATE	0.0000	0.0000	0.0000	0.0514	0.0010
EURO 1 YEAR OIS - MIDDLE RATE	0.0000	0.0000	0.0000	0.0001	0.0957
ECU EURO-ECU 5 MTH - MIDDLE RATE	0.4303	0.2631	0.0041	0.8175	0.2949
ECU EURO-ECU 1 YEAR - MIDDLE RATE	0.0000	0.0510	0.0000	0.3933	0.2008
IBA EUR IBK. LIBOR 1M DELAYED - OFFERED RATE	0.3617	0.0021	0.4762	0.1522	0.1413
IBA EUR IBK. LIBOR 6M DELAYED - OFFERED RATE	0.0018	0.0067	0.0000	0.1316	0.1280
IBA EUR IBK. LIBOR O/N DELAYED - OFFERED RATE	0.0843	0.4895	0.8353	0.2588	0.0795
(TR) EURO VS EURIBOR 6M IR SWAP 5Y - MIDDLE RATE	0.0000	0.0000	0	0.0000	0.2583
INEU03(CM01)	0.0000	0.0000	0.0000	0.3865	0.8977
INEU03(CM10)	0.0000	0.0000	0	0.0000	0.0455
INEU03(CM05)	0.0000	0.0000	0	0.0000	0.1951
INEU03(CM02)	0.0000	0.0000	0.0000	0.0068	0.3690

Appendix C: Rolling window results

A) 1% CAViaR

Table 1. Results for the First Model 1% VaR

	Coeff estimated	Signif Coeff	Mean	Std. Dev.	Min	Max	Positive Coeff	Negative Coeff
Constant Fund1	484	1	-0,12819	0	-0,12819	-0,12819	0	1
Constant Fund2	484	0	0	0	0	0	0	0
Constant Fund3	484	30	-1,08097	1,069567	-3,77116	-0,2666	0	30
AutoregQuant Fund1	484	483	0,889657	0,032772	0,787419	0,964299	483	0
AutoregQuant Fund2	484	413	0,843087	0,072131	0,58581	0,9415	413	0
AutoregQuant Fund3	484	481	0,835215	0,174716	-0,73484	0,970482	477	4
Abs Comp Fund1	484	29	-0,19793	0,024271	-0,22162	-0,08435	0	29
Abs Comp Fund2	484	26	-0,201	0,016426	-0,25367	-0,16853	0	26
Abs Comp Fund3	484	77	-0,32483	0,107844	-0,53703	-0,1455	0	77

Table 2. Results for the Second Model 1% VaR

	Coeff estimated	Signif Coeff	Mean	Std. Dev.	Min	Max	Positive Coeff	Negative Coeff
Constant Fund1	484	313	-0,12938	0,038384	-0,45223	-0,01609	0	313
Constant Fund2	484	447	-0,32245	0,184299	-0,92875	-0,08907	0	447
Constant Fund3	484	303	-0,28269	0,337705	-4,00572	-0,05276	0	303
AutoregQuant Fund1	484	482	0,891534	0,039453	0,622989	0,998767	482	0
AutoregQuant Fund2	484	481	0,755364	0,137107	0,354382	0,964309	481	0
AutoregQuant Fund3	484	474	0,847915	0,109879	0,506282	0,9836	474	0
Abs Comp Fund1	484	142	-0,20928	0,036314	-0,30152	0,018615	1	141
Abs Comp Fund2	484	151	-0,37496	0,083379	-0,51931	-0,24727	0	151
Abs Comp Fund3	484	14	-0,29572	0,244762	-0,8819	-0,16915	0	14
Benchmark Fund1	484	397	0,270008	0,050587	0,082794	0,457363	397	0
Benchmark Fund2	484	468	0,394812	0,122882	0,191148	0,613193	468	0
Benchmark Fund3	484	372	0,30001	0,102338	0,065542	0,688683	372	0

Table 3. Results for the Third Model 1% VaR

	Coeff estimated	Signif Coeff	Mean	Std. Dev.	Min	Max	Positive Coeff	Negative Coeff
Constant Fund1	484	105	-0,17289	0,310252	-2,78532	-0,05016	0	105
Constant Fund2	484	390	-0,4122	0,300045	-1,41047	-0,03167	0	390
Constant Fund3	484	356	-0,60913	0,761122	-4,04113	-0,08091	0	356
AutoregQuant Fund1	484	479	0,870586	0,086259	-0,44177	0,970082	478	1
AutoregQuant Fund2	484	434	0,758269	0,183545	0,254116	1,000163	434	0
AutoregQuant Fund3	484	394	0,844517	0,121714	0,298894	0,979254	394	0
Abs Comp Fund1	484	46	-0,2464	0,118867	-0,4147	0,394522	1	45
Abs Comp Fund2	484	134	-0,63471	0,187605	-1,12386	0,057552	1	133
Abs Comp Fund3	484	88	-0,54367	0,215752	-0,94995	-0,10346	0	88
Benchmark Fund1	484	68	0,327933	0,095326	0,143472	0,516445	68	0
Benchmark Fund2	484	205	0,483252	0,160316	0,094452	0,680628	205	0
Benchmark Fund3	484	143	0,581262	0,365042	-0,09691	1,091565	140	3
P.C. 1 Fund 1	484	27	0,086292	0,050953	-0,1224	0,142767	26	1
P.C. 1 Fund 2	484	87	0,122978	0,055698	0,06284	0,214601	87	0
P.C. 1 Fund 3	484	157	-0,01399	0,048162	-0,06699	0,077506	39	118
P.C. 2 Fund 1	484	236	-0,124	0,022621	-0,1601	-0,05494	0	236
P.C. 2 Fund 2	484	117	-0,17695	0,082547	-0,31347	-0,06731	0	117
P.C. 2 Fund 3	484	190	0,094858	0,025842	0,047671	0,144071	190	0
P.C. 3 Fund 1	484	44	0,121572	0,020167	0,052325	0,143241	44	0
P.C. 3 Fund 2	484	77	0,126429	0,045753	-0,09242	0,215072	76	1
P.C. 3 Fund 3	484	74	-0,00502	0,096554	-0,07679	0,226624	13	61
P.C. 4 Fund 1	484	119	0,318252	0,096354	0,12644	0,562598	119	0
P.C. 4 Fund 2	484	191	0,346089	0,072158	0,213212	0,570215	191	0
P.C. 4 Fund 3	484	55	-0,13848	0,01959	-0,16905	-0,09866	0	55
P.C. 5 Fund 1	484	70	-0,22352	0,046552	-0,32626	-0,1354	0	70
P.C. 5 Fund 2	484	83	-0,25769	0,061304	-0,32839	-0,1054	0	83
P.C. 5 Fund 3	484	80	-0,07824	0,09434	-0,21521	0,179877	10	70

B) 5% CAViaR

Table 4. Results for the First Model 5% VaR

	Coeff estimated	Signif Coeff	Mean	Std. Dev.	Min	Max	Positive Coeff	Negative Coeff
Constant Fund1	484	0	0	0	0	0	0	0
Constant Fund2	484	72	-0,26225	0,481896	-2,25694	-0,11621	0	72
Constant Fund3	484	5	-1,10775	1,284739	-2,51718	-0,16943	0	5
AutoregQuant Fund1	484	480	0,890847	0,021701	0,834075	1,003309	480	0
AutoregQuant Fund2	484	436	0,865633	0,106599	-1,00479	0,993903	435	1
AutoregQuant Fund3	484	480	0,888351	0,031543	0,829455	1,00737	480	0
Abs Comp Fund1	484	183	-0,19426	0,028143	-0,26409	-0,10982	0	183
Abs Comp Fund2	484	276	-0,1833	0,019518	-0,22162	-0,14838	0	276
Abs Comp Fund3	484	237	-0,21361	0,051574	-0,29109	-0,10417	0	237

Table 5. Results for the Second Model 5% VaR

	Coeff estimated	Signif Coeff	Mean	Std. Dev.	Min	Max	Positive Coeff	Negative Coeff
Constant Fund1	484	323	-0,06752	0,078685	-0,80403	-0,03178	0	323
Constant Fund2	484	483	-0,44551	0,199986	-0,88616	-0,07729	0	483
Constant Fund3	484	444	-0,35745	0,463877	-1,4824	-0,03937	0	444
AutoregQuant Fund1	484	483	0,905111	0,064109	0,317074	0,945575	483	0
AutoregQuant Fund2	484	450	0,55583	0,162925	0,19443	0,935855	450	0
AutoregQuant Fund3	484	404	0,842928	0,125146	0,387922	0,938401	404	0
Abs Comp Fund1	484	25	-0,15798	0,096116	-0,43147	-0,06913	0	25
Abs Comp Fund2	484	336	-0,33721	0,044483	-0,50173	-0,21163	0	336
Abs Comp Fund3	484	87	-0,29041	0,081176	-0,45832	-0,10157	0	87
Benchmark Fund1	484	482	0,141066	0,041781	0,089137	0,423618	482	0
Benchmark Fund2	484	483	0,450297	0,084127	0,236671	0,622527	483	0
Benchmark Fund3	484	480	0,251127	0,108662	0,116277	0,54975	480	0

Table 6. Results for the Third Model 5% VaR

	Coeff estimated	Signif Coeff	Mean	Std. Dev.	Min	Max	Positive Coeff	Negative Coeff
Constant Fund1	484	209	-0,42433	0,309367	-1,83963	-0,03829	0	209
Constant Fund2	484	474	-0,55983	0,161476	-0,8608	-0,09247	0	474
Constant Fund3	484	411	-0,45636	0,372968	-1,30127	-0,05685	0	411
AutoregQuant Fund1	484	443	0,805913	0,208035	-0,63799	0,979598	442	1
AutoregQuant Fund2	484	453	0,421842	0,142345	0,180519	0,866951	453	0
AutoregQuant Fund3	484	406	0,706191	0,216695	0,264747	0,959161	406	0
Abs Comp Fund1	484	163	-0,33805	0,130751	-0,54587	-0,08387	0	163
Abs Comp Fund2	484	359	-0,35433	0,067049	-0,57622	-0,12538	0	359
Abs Comp Fund3	484	111	-0,27695	0,07574	-0,53073	-0,16257	0	111
Benchmark Fund1	484	181	0,409449	0,155824	0,112472	0,620183	181	0
Benchmark Fund2	484	444	0,410027	0,104335	0,144161	0,699568	444	0
Benchmark Fund3	484	166	0,380274	0,084761	0,125141	0,5548	166	0
P.C. 1 Fund 1	484	51	0,051176	0,017429	0,030223	0,086651	51	0
P.C. 1 Fund 2	484	138	0,143947	0,032874	0,063898	0,221757	138	0
P.C. 1 Fund 3	484	282	-0,03752	0,011049	-0,06431	-0,02101	0	282
P.C. 2 Fund 1	484	286	-0,09223	0,011967	-0,13224	-0,06055	0	286
P.C. 2 Fund 2	484	90	-0,19358	0,077594	-0,3263	-0,0465	0	90
P.C. 2 Fund 3	484	283	0,070932	0,010002	0,031851	0,104719	283	0
P.C. 3 Fund 1	484	144	0,103186	0,01111	0,061914	0,143099	144	0
P.C. 3 Fund 2	484	105	0,174571	0,041966	0,069351	0,257052	105	0
P.C. 3 Fund 3	484	30	-0,0544	0,004187	-0,06103	-0,04253	0	30
P.C. 4 Fund 1	484	281	0,245902	0,053538	0,150036	0,43874	281	0
P.C. 4 Fund 2	484	367	0,347073	0,078434	0,154956	0,469969	367	0
P.C. 4 Fund 3	484	11	-0,04619	0,068696	-0,09404	0,102028	2	9
P.C. 5 Fund 1	484	113	-0,14304	0,028695	-0,23305	-0,09044	0	113
P.C. 5 Fund 2	484	253	-0,1988	0,042455	-0,30968	-0,11387	0	253
P.C. 5 Fund 3	484	46	-0,14183	0,060745	-0,20574	0,185136	1	45

REFERENCES

Koenker, R. (2005). *Quantile regression*. Cambridge: Cambridge University Press.

Mosteller, F. and Tukey, J. (1977). *Data analysis and regression*. Reading, Mass.: Addison-Wesley Pub. Co.

Montenegro, A. (2001). *On sample size and precision in ordinary least squares*. Journal of Applied Statistics, 28(5), pp.603-605.

Baum, C. (2013). *Quantile Regression*.

Koenker, R. and Hallock, K. (2001). *Quantile Regression*. Journal of Economic Perspectives, 15(4), pp.143–156.

Koenker, R. and Hallock, K. (2000). *Quantile Regression an Introduction*. 1st ed. University of Illinois at Urbana-Champaign.

Chen, C. (n.d.*). An Introduction to Quantile Regression and the QUANTREG Procedure*. 1st ed. SAS Institute Inc., Cary, NC.

He, X. and Wei, Y. (2005). *Tutorial on Quantile Regression*.

Allen, D. and Singh, A. (2009). *Minimizing Loss at Times of Financial Crisis: Quantile Regression as a Tool for Portfolio Investment Decisions*. 1st ed. School of Accounting, Finance and Economics, Cowan University.

Stefanini, F. (2009). *Corso di Risk Management - Value at Risk*.

Meneguzzo, D. and Vecchiato, W. (n.d.). *Improvements on Value at Risk measures by combining conditional autoregressive and extreme value approaches*. 1st ed.

Engle, R. and Manganelli, S. (2002*). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles*.

Allen, D. and Singh, A. (2010). *CAViaR and the Australian Stock Markets: An Appetiser*. SSRN Electronic Journal.

Engle, R. and Manganelli, S. (2004*). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles*. Journal of Business & Economic Statistics, 22(4), pp.367-381.

Kouretas, G. and Zarangas, L. (2005). *Conditional Autoregressive Value at Risk by Regression Quantiles: Estimating market risk for major stock markets*.

Kidd, D. (n.d.). *Value at Risk and Conditional Value at Risk: A Comparison*. CFA Institute.

Furlan, S. (2004). *Il Value at Risk per la gestione del rischio di mercato: metodi di calcolo e procedure di backtesting*. Università degli studi di Padova.

Sarykalin, S., Serraino, G. and Kalinchenko, K. (n.d.). *VaR vs CVaR in Risk Management and Optimization*.

White, H., Kim, T. and Manganelli, S. (2015). *VAR for VaR: Measuring tail dependence using multivariate regression quantiles*. Journal of Econometrics, 187(1), pp.169-188.

Global Optimization Toolbox User's Guide. (2014). The MathWorks, Inc.

Principal Component Analysis. (n.d.). SAS Institute. Available at: http://support.sas.com/publishing/pubcat/chaps/55129.pdf [Accessed 8 Dec. 2015].

Johnson, R. and Wichern, D. (1998). *Applied multivariate statistical analysis*. 4th ed. Englewood Cliffs, N.J.: Prentice-Hall, pp.458-512.

Multiple Authors (n.d.). *Rischio e rendimento nella gestione del risparmio: misura, controllo, attribuzione*. Assogestioni, pp.1-145.

En.wikibooks.org, (n.d.). Statistics/Numerical Methods/Quantile Regression - Wikibooks, open books for an open world. [online] Available at: https://en.wikibooks.org/wiki/Statistics/Numerical_Methods/Quantile_Regression [Accessed 10 Nov. 2015].

Wikipedia, (2010). *Quantile regression*. [online] Available at: https://en.wikipedia.org/wiki/Quantile_regression [Accessed 10 Nov. 2015].

Wikipedia, (n.d.). *Genetic algorithm*. [online] Available at: https://en.wikipedia.org/wiki/Genetic_algorithm [Accessed 1 Dec. 2015].

Morningstar.it, (2015). Capital Group European Growth and Income Fund (LU...|LU0342012266. [online] Available at: http://www.morningstar.it/it/funds/snapshot/snapshot.aspx?id=F00000SZFU [Accessed 09 Dec. 2015].

It.mathworks.com, (n.d.). *Principal Component Analysis (PCA)* - MATLAB & Simulink. [online] Available at: http://it.mathworks.com/help/stats/principal-component-analysis-pca.html [Accessed 15 Dec. 2015].

YouTube, (2010). *What is Principal Component Analysis (PCA)?*. [online] Available at: https://www.youtube.com/watch?v=f9mZ8dSrVJA [Accessed 15 Dec. 2015].

Support.minitab.com, (n.d.). *What is principal components analysis? - Minitab*. [online] Available at: http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/multivariate/principal-components-and-factor-analysis/what-is-pca/ [Accessed 17 Dec. 2015].

Simonemanganelli.org, (n.d.). Simone Manganelli | Research. [online] Available at: http://www.simonemanganelli.org/Simone/Research.html [Accessed 18 Dec. 2015].

Lexicon.ft.com, (2012). *Value At Risk - Var Definition from Financial Times Lexicon*. [online] Available at: http://lexicon.ft.com/Term?term=value-at-risk-_-VaR [Accessed 20 Dec. 2015].

Morningstar.it, (2016). Henderson Horizon Fund - Pan European Equity Fund...|LU0138821268. [online] Available at: http://www.morningstar.it/it/funds/snapshot/snapshot.aspx?id=F0GBR04NWX [Accessed 14 Jan. 2016].

Morningstar.it, (2016). JPM Funds - Europe Dynamic Fund A (dist) - EUR|LU0119062650. [online] Available at: http://www.morningstar.it/it/funds/snapshot/snapshot.aspx?id=F0GBR04CLM [Accessed 14 Jan. 2016].

Howard D. Bondell, H. (2010). *Noncrossing quantile regression curve estimation. Biometrika*, [online] 97(4), p.825. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3371721/ [Accessed 1 Feb. 2016].