Università degli Studi di Milano-Bicocca

Dipartimento di Statistica e Metodi Quantitativi

Dottorato di ricerca in Matematica per l'Analisi

dei Mercati Finanziari

XXVI Ciclo

Risk attribution and semi-heavy tailed distributions

Tesi di Dottorato di

Edit Rroji

matricola 076099

Relatore : Prof. Marco Marchioro

Tutor : Prof. Emanuela Rosazza Gianin

A.A. 2013/2014

Contents

Thank you.

Introduction

After the recent financial crisis there is much more emphasis on the sources of risk rather than just on the levels. The importance of attributing risk to each portfolio component can be deduced even from the fact that in the last few years a new principle arose in finance: the Risk Parity. It is an approach in portfolio management which focuses on allocation of risk rather than on the capital allocation (see Denis *et al.* [2011] for further details).
Value at risk (see RiskMetrics [1996]) and Expected Shortfall (see Tasche [2002]) have emerged as industry standards for measuring downside risk. Non-parametric approaches have gained the consensus of the practitioners since they are easier to implement and only the information in the return series is used. However, usually non-parametric methods imply more uncertain estimates for the risk measures considered as shown for example in Aussenegg and Miazhynskaia [2006]. Although risk decomposition in a parametric context is not straightforward, Boudt *et al.* [2009] show that the use of modified versions of VaR and ES based on asymptotic expansions simplifies the problem.
In this thesis we discuss the problem of risk attribution in a multifactor context using non-parametric approaches but we also introduce a new distribution for modeling returns. The risk measures considered are homogeneous since we exploit the Euler rule as in Tasche [1999]. Particular attention is given to the problem of attributing risk to user defined factors since the existing literature is limited when compared to other research arguments but of practical relevance. We point out the problems encountered during the analysis and present some methodologies that can be useful in practice. Each chapter combines both theoretical and practical issues.
The main original contributions can be found in Chapter 3 and Chapter 4 respectively in the introduction of a methodology for identifying idiosyncratic risk in presence of custom factors and the introduction of a new distribution named Mixed Tempered Stable.
The thesis is organized as follows. In Chapter 1 we review the existing literature on linear multi-factor models. Starting from the usual Ordinary Least Squares regression we move on to a more sophisticated method for factor exposure estimation, the Kalman Filter (see Kalman [1960]). Principal Component Analysis (see Jolliffe [2002]) is applied for the identification of statistical factors in order to reduce dimensionality. At the end of the chapter a comparison with another mathe-

matical procedure, the Independent Component Analysis (see Hyvarinen [1999]), is given.
In Chapter 2 we review the main characteristics of two well known homogeneous risk measures : the Value At Risk (VaR) and the Expected Shortfall (ES). We discuss the problem of defining a unique risk measure from the regulator point of view and the mathematical properties required. The expectiles introduced in Newey and Powell [1987] as the solution of an Asymmetric Least Squares (ALS) regression have been considered only recently in the context of risk measurement. Coherence (see Artzner *et al.* [1999], Bellini *et al.* [2013]) and elicitability (see Gneiting [2011]) are two desirable properties that are both satisfied by the expectile based risk measure. The main difficulty in its use is the identification of the parameter that controls the tail weight in the ALS regression which in Kuan *et al.* [2009a] is interpreted as the level of prudentiality. From the quantile - expectile relation derived in Jones [1994], we find the implicit values for the weight parameter in datasets coming from equity and credit risk markets. The confidence levels used are those usually fixed by the regulators in the Value at Risk computation. The estimated parameter values are then used for comparing risk attribution results for the three considered risk measures.
Chapter 3 is devoted to the study of risk decomposition models for arbitrary factors chosen directly from the investors. We describe the approach proposed in Meucci [2007] and extend the analysis by considering independent factors obtained through the Independent Component Analysis Hyvarinen [1999] since the methodology can be applied to any linear transformation of the original risk factors that generates new uncorrelated variables. Starting from the work of Menchero and Poduri [2008] we derive a model for attributing risk to generic factors. The introduction of new factors, orthogonal to those considered by the investor, enables us to identify the idiosyncratic term and its contribution to risk. This result can be very useful in risk management because it is simple to implement and each component has a straightforward interpretation.
We propose a new distribution, named Mixed Tempered Stable (MixedTS), in Chapter 4. It is built as a Normal Variance Mean Mixture (NVMM) where the Normal assumption is substituted with the standardized Tempered Stable distribution. Assuming that the mixing random variable follows a Gamma distribution, the proposed model has the Variance Gamma (see Madan and Seneta [1990b]), the Tempered Stable (see Kim *et al.* [2008]) and the Geometric Stable distribution (see Kozubowski *et al.* [2011]) as special cases for particular parameters choice. The moment generating function exists and we give analytical formulas for the first four moments. The MixedTS, having as special cases some distributions widely applied in financial literature, is expected to be flexible in capturing different tail behaviors. By construction, no more than one independent component obtained through an ICA analysis can be Gaussian though it is interesting to see how the MixedTS fits the different components. Risk measures for the MixedTS are computed using numerical integration and the Saddlepoint approximation (Lugannani and Rice [1980]) for the tails.

Chapter 1

Linear factor models

In this chapter we review the literature on linear multifactor models for describing returns and discuss the problems arising in their use. The return decomposition is the starting point for the risk attribution process that will be discussed in the next chapters.

The Ordinary Least Squares regression is the simplest way of facing these kind of problems. It is well known that it considers the entire estimation period for computing factor exposures. By using the Kalman Filter (Kalman [1960]) we have a model where factor exposures can vary daily while maintaining the information obtained from the past. Kalman Filter is based on a state space model where the state variables are the factor exposures. The implementation is not easy and part of the data is needed for the algorithm stability. Both OLS and Kalman Filter can be applied when factor time series are given.

Principal Component Analysis Jolliffe [2002] generates statistical factors that are the projections of the original data on the eigenvectors associated to the dataset covariance matrix. In finance this technique is particularly useful for dimension reduction purposes. In a recent paper Longstaff *et al.* [2011] study the nature of sovereign credit risk and find that the majority of sovereign credit risk can be linked to global factors. They find that sovereign credit spreads are more related to the stock indexes than they to local economic measures. In the same spirit we apply the PCA analysis to credit spreads of European companies grouped by their rating. The dependence on different factors can suggest that the structure of each group is different.

Principal Component Analysis has been widely applied although the gaussianity hypothesis for the components is not in line with the stylized facts observed in financial markets. Independent Component Analysis (ICA) can be a valid alternative since it yields independent factors by maximizing the non-gaussianity of the components. Unfortunately, in the ICA analysis there is no single rule for ordering the components while the explained variance is a criterion for ordering the principal components. This fact justifies the difficulties of using ICA in dimension reduction problems in finance.

1.1 Multifactor models in finance

The simplest type of models are linear multifactor models since the relation between the variable of interest and the factors from which it is influenced is linear. Let us consider the portfolio return π^t for the time interval $(t-1;t]$ defined as :

$$\pi^t = \sum_{i=1}^{n} w_i^t r_i^t \tag{1.1}$$

where w_i^t is the weight of the i-th asset in the portfolio. The portfolio is composed by n assets with the i-th asset log-return [1], given its price P_i^t , for the period $(t-1;t]$ is :

$$r_i^t = \ln\left(\frac{P_i^t}{P_i^{t-1}}\right) \tag{1.2}$$

We suppose that the investment decision is taken at time $t-1$ and the holding period is the time interval $(t-1;t]$. We assume a linear relation with a finite number of factors F^t :

$$r_i^t = \alpha_i^t + \sum_{j=1}^{n'} \beta_{ij}^t F_{ij}^t + \epsilon_i^t \tag{1.3}$$

with α^t being a constant and ϵ^t the residual term. The risk factors F[2] are financial variables like log returns on equities, bonds or indexes while from quantities like interest rates and spreads risk factors are built using absolute variations. The literature on this argument is rich and in continuous evolution. The main contributions in chronological order are :
CAPM one factor model (see Sharpe [1964], Lintner [1965] and Black [1972]).
It assumes a linear relationship between the expected return of a risky asset and its β:

$$E[r_i] = r_f + \beta_i(E[r_{mkt}] - r_f) \tag{1.4}$$

where r_{mkt} is the return of the well diversified market porfolio and r_f the risk free rate. The model states that β_i defined as :

$$\beta_i = \frac{Cov(r_i, r_{mkt})}{Var(r_{mkt})} \tag{1.5}$$

drives the expected returns of asset `i` since it measures the part of the asset's statistical variance that cannot be removed by the diversification provided in the market portfolio.
Fama and French three factor model can be defined in the form (see Fama and French [1992] and

[1]We use log returns instead of simple returns since they are additive, i.e longer period returns are obtained by summation of short period returns.
[2]From now on we will drop the time index and we will consider only one period returns

Fama and French [1995]):

$$E[r_i] = \beta_i{}^{mkt}E[r_{mkt}] + \beta_i{}^{SMB}E[r_{SMB}] + \beta_i{}^{HML}E[r_{HML}] \tag{1.6}$$

They assume that all market returns can be explained by three factors:the exposure to the portfolio market, the exposure to value stocks (HML: High [book-to-market ratio] Minus Low), and the exposure to small stocks (SMB: Small [market capitalization]) where r_{SMB} measures the excess returns of small caps over big caps and r_{HML} the excess returns of value stocks over growth stocks. Carhart [1997] four factor model considers the momentum in addition to the three factor model:

$$E[r_i] = \beta_i{}^{mkt}E[r_{mkt}] + \beta_i{}^{SMB}E[r_{SMB}] + \beta_i{}^{HML}E[r_{HML}] + \beta_i{}^{Mom}E[r_{Mom}] \tag{1.7}$$

The recent literature considers a higher number of factors. For example the Barra Integrated Model (BIM) considers more than 200 factors (see Shepard [2011]). It identifies first local risk factors and develop local models for different asset classes. These are then combined into a single risk model to capture cross-market correlations. The resulting model provides a structure for detailed risk decomposition of multi-asset class strategies.

Factors can have an economic interpretation, or can be statistical. In particular, in Straumann and Garidi [2011] they distinguish between six type of factors: macroeconomic, market, sector/industry, fundamental, technical and statistical.
Macroeconomic factors can be for example change in the growth rate of GDP or CPI, differences between returns on high yield and government bond indexes etc. Since most of the macroeconomic variables are released quarterly and risk management is performed daily, their use presents some problems. Returns on indexes like for example S&P500 or FTSEMib are market factors. Morgan Stanley Capital International (MSCI) releases daily estimates on GICS (Global Industry Classification Standard) indexes and the corresponding returns are the so called sector/industry factors. In 2004, MSCI acquired Barra Inc. which uses these factors especially in the Barra Integrated Equity Model.
Firm specific quantities like price to book ratios and market capitalization are considered as fundamental factors while typical examples for technical factors are measures of liquidity. Statistical factors are different since they are not directly observable in the market but extracted via mathematical procedures.
These models have been used for return attribution in Brinson [1985] and risk contribution in Meucci [2007], Menchero and Poduri [2008] and Marchioro and Borrello [2013]. After the decision of the type of factors to consider, it remains to fix the number of factors to include in the model. Except for the cases when one is fully confident about the factors to include, the decision should be taken after considering some statistical tests. One simple way is to proceed step by step, in the sense that when we add an additional factor we compare the likelihood of the model with n factors to that of the n+1 factors model. Under the i.i.d assumption, for large time-series samples, the

ratio of the log likelihoods has an approximate chi-squared distribution. In particular, Connor and Korajczyk (1993) derive a test for the number of factors that is robust and is based on the decline in average idiosyncratic variance as additional factors are added.
In matrix notations, the linear factor model is :

$$r = \alpha + \beta F + \epsilon \tag{1.8}$$

where r, a and ϵ are $1 \times T$ vectors. If we consider K factors β is a $1 \times K$ vector, while F is a $K \times T$ matrix. Meucci [2010] puts the conditions for a model to be in the systematic-plus-idiosyncratic form : the residuals must be uncorrelated with each other and with the factors. In this paper he shows that the multifactor models used in finance are never of this form.
In models considering statistical risk factors both factors F and exposures β must be estimated but no constraints on the residual uncorrelation are posed. For risk factors directly observed in the market, matrix F is fully known and what remains to estimate is only the factor exposures β and the vector of constants α. Models of this type are coined as time-series models.
The Ordinary Least Squares approach is obtained by considering time-series models for which factor exposures are the solution of an unconstrained minimization problem. In mathematical notations, the factor exposure estimates $\hat{\beta}$ resulting from the OLS approach are the solution of the problem :

$$\hat{\beta} = argmin_{\beta} \, \|r - \beta F\| \tag{1.9}$$

where $\|.\|$ is the L^2-norm. Geometrically, it projects the returns in r on the linear space spanned by the factors F. The minimization problem give the estimate of the factor exposures in matrix notations:

$$\hat{\beta} = (FF')^{-1} Fr \tag{1.10}$$

Particular attention deserves the decision of the time frame to consider for factor exposure estimation. In fig 1.2 we show the results of the OLS regression for four fund indexes vs 17 factors. The first fund considered is the Vanguard Asset allocation Fund which tracks the performance of its benchmark index, the $S\&P500$ index. The second fund is the Vanguard Balanced Fund which tracks the performance of the overall US stock market index using 60 % of its capital and with the remaining tracks the performance of a market-weighted bond index. Allianz RCM Global Water and First Eagle Gold funds invest mainly in assets related to commodities. Indeed, the first invests on equity securities related to companies that have direct or indirect exposures to water related activities while the last fund invests 80 % of its assets in gold or related securities.
The factors used for explaining the returns for each fund are based on different asset classes. Two commodity indexes: Oil and Gold ; three interest rate based risk factors: 3m Libor, 10y and 30 swap rates ; the exchange rate EurUsd, the implied volatility index VIX and the 10 GICS sector indexes.

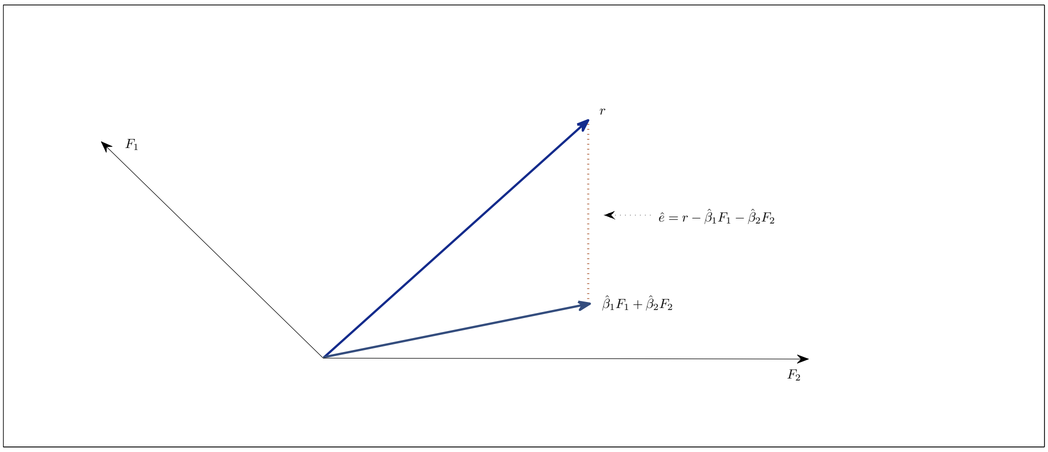


Figure 1.1: Geometric interpretation of the OLS regression: projection of returns in r on the linear space spanned by the regressors F_1 and F_2. The factor exposures estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ minimize the residual $\hat{e}$.

The evaluation date is 12/10/2012. The dataset is composed by daily log returns while the time frame estimation period changes. We consider respectively 1.5, 2, 2.5 and 3 years as estimation periods and compare the results. In general, factor exposures tend to remain stable but this is not always the case. Consider for example the Vanguard Balanced Fund exposure to the 3m Libor : choosing to use the last 2 years data instead of 2.5 years doubles the factor exposure.
The sign of the exposures remains the same since, depending on the target of each fund, the relation between each fund and the risk factor has an economic interpretation that is translated in a numeric value. For example, Allianz RCM Global Water fund seems to have a greater exposure to the Industrial GICS index. This is coherent with the idea that water influences the industrial sector. As we expected, the First Eagle Gold fund returns are highly influenced from the Gold returns. The OLS regression is very useful in giving an intuition of the factors that we must consider for a correct return/risk attribution model but we have to move on to more advanced instruments for computing factor exposures that are updated at least daily.

Vanguard Asset Allocation Inv

Tracks the performance of the SP500 index

	WTI Oil	Gold	US003M	SWAP10	SWAP30	EURUSD	VIX	INFT	FINL	CONS	HLTH	ENRS	COND	INDU	MATR	UTIL	TELS
1,5	2,9E-04	0,004	-0,002	0,010	-0,008	-0,009	1,5E-04	0,186	0,137	0,134	0,096	0,118	0,114	0,111	0,043	0,042	0,024
2	3,8E-04	0,004	0,001	0,006	-0,005	-0,007	-1,4E-04	0,185	0,143	0,127	0,098	0,120	0,112	0,112	0,038	0,042	0,024
2,5	-9,5E-05	0,005	0,001	0,007	-0,006	-0,002	-3,2E-04	0,182	0,147	0,119	0,103	0,119	0,111	0,112	0,037	0,041	0,022
3	-1,3E-04	0,004	0,001	0,005	-0,004	-0,001	-3,6E-04	0,185	0,148	0,120	0,106	0,117	0,111	0,111	0,036	0,039	0,023

Vanguard Balanced Index

With 60% of the capital tracks the performance of the overall US stock market and with 40% the performance of a market- weighted bond index.

	WTI Oil	Gold	US003M	SWAP10	SWAP30	EURUSD	VIX	INFT	FINL	CONS	HLTH	ENRS	COND	INDU	MATR	UTIL	TELS
1,5	0,007	0,030	0,118	-0,050	0,025	0,108	0,018	0,158	0,065	0,136	0,079	0,036	0,136	0,032	0,027	0,027	0,047
2	0,006	0,033	0,130	-0,045	0,017	0,090	0,018	0,159	0,074	0,133	0,076	0,042	0,140	0,032	0,027	0,030	0,037
2,5	0,010	0,056	0,049	-0,033	0,003	0,098	0,022	0,157	0,069	0,149	0,095	0,031	0,144	0,045	0,018	0,051	0,037
3	0,009	0,049	0,056	-0,037	0,008	0,097	0,022	0,159	0,076	0,157	0,104	0,032	0,141	0,048	0,007	0,042	0,028

Allianz RCM Global Water

Seek the long-term capital appreciation. Invest in equity securities of companies which have direct or indirect exposure to water related activities.

	WTI Oil	Gold	US003M	SWAP10	SWAP30	EURUSD	VIX	INFT	FINL	CONS	HLTH	ENRS	COND	INDU	MATR	UTIL	TELS
1,5	0,025	0,046	-0,103	0,006	-0,004	0,213	-0,022	0,085	0,081	-0,003	0,083	-0,013	0,010	0,325	0,086	0,170	0,014
2	0,020	0,051	-0,110	0,019	-0,021	0,210	-0,018	0,071	0,065	-0,005	0,090	0,016	0,023	0,322	0,099	0,134	0,036
2,5	0,031	0,037	-0,066	0,006	-0,007	0,221	-0,018	0,072	0,059	0,001	0,084	0,015	0,022	0,350	0,079	0,067	0,047
3	0,032	0,024	-0,059	0,001	0,003	0,219	-0,017	0,085	0,039	0,017	0,084	0,017	0,016	0,356	0,074	0,054	0,052

First Eagle Gold

It invests 80% of its assets in gold or related securities.

	WTI Oil	Gold	US003M	SWAP10	SWAP30	EURUSD	VIX	INFT	FINL	CONS	HLTH	ENRS	COND	INDU	MATR	UTIL	TELS
1,5	0,041	0,898	-0,071	0,012	0,030	0,012	-0,012	0,158	-0,073	-0,265	-0,109	0,090	-0,005	-0,236	0,508	0,059	-0,035
2	0,033	0,945	-0,122	0,083	-0,042	0,048	-0,013	0,141	-0,047	-0,286	-0,133	0,085	0,084	-0,295	0,473	0,060	-0,012
2,5	0,046	0,941	-0,072	0,075	-0,049	-0,003	-0,017	0,119	-0,054	-0,225	-0,101	0,080	0,047	-0,197	0,415	0,017	-0,023
3	0,033	0,954	-0,075	0,077	-0,054	0,002	-0,018	0,069	-0,071	-0,186	-0,110	0,085	0,032	-0,129	0,445	-0,011	-0,043

Figure 1.2: We consider different funds and through a simple linear regression obtain the exposures to the selected risk factors. To emphasize the importance of estimation window selection we report results for four estmation windows: 1.5, 2, 2.5 and 3 years.

1.1.1 Kalman Filter

An alternative method for estimating factor exposures that does not require any decision on the time frame to consider is the Kalman Filtering (KF) introduced in Kalman [1960]. It is based on a state space model composed from two equations namely state and observation equation. The Kalman filter is a recursive algorithm that computes estimates for the unobserved variables at time t, based on the available information up to that date. The filter structure and the update after receiving a new information give more reliable results than a simple Ordinary Least Squares regression. Here we outline how the Kalman Filter can be used for estimation.
Let r_t be a $n \times 1$ be the vector of n asset returns at time t. Suppose that each asset return depends linearly on p factors that are the rows of the deterministic $n \times p$ matrix H_t. The state variables are the factor exposures in the $p \times 1$ vector β_t. A general linear state space model is of the form :

$$r_t = H_t \beta_t + \epsilon_t \tag{1.11}$$

$$\beta_t = A_t \beta_{t-1} + v_t \tag{1.12}$$

where $\epsilon_t \sim N(0, R_t)$ and $v_t \sim N(0, Q_t)$. We assume ϵ_t and v_t to be mutually and serially uncorrelated, meaning that they don't depend on each other and on their past values. We will refer to equation 1.11 as the observation (or measurement) equation, while equation 1.12 is the state equation which describes the dynamic of the unobserved state variable β_t. If $A_t = I$, the distribution of each factor loading in β is a random walk. In financial applications it is reasonable to suppose A_t different from the identity matrix in order to model autocorrelated factor exposures because we observe persistence in financial market returns. In the case of a time homogeneous system, matrices A_t,R_t and Q_t don't depend on time. We assume stationarity for the shocks and for the autocorrelation structure in the state varibles, i.e we consider an homogeneous model.

Once the model is put in the state space form, the Kalman Filter is used to compute optimal forecasts of the normally distributed state space vector β_t, at a given time t, based on the information we have up to the previous time $t-1$. The purpose of the filter is to update the estimate for the state vector as soon as we observe r_t. Consider the information set I_t generated by the return observations and by the estimates of the state space variable up to time t.

$$I_{t-1} = \left\{ r_{t-1}, ..., r_0, \hat{\beta}_{t-1}, ..., \hat{\beta}_0 \right\} \tag{1.13}$$

It contains the past return observations $r_{t-1}, ..., r_0$ and the past estimates $\hat{\beta}_{t-1}, ..., \hat{\beta}_0$. The a priori estimates of the state variables are the best estimates using the information contained in I_{t-1}:

$$\hat{\beta}_t^* = E\left[\beta_t \left| I_{t-1} \right.\right] \tag{1.14}$$

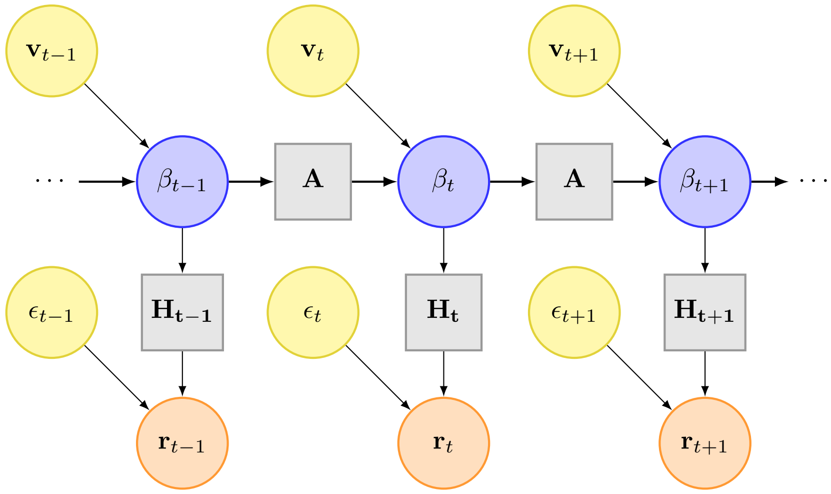


Figure 1.3: In this figure we describe the Kalman Filter model. Equation 1.12 formulates the dynamics of the state variables in β_t while the relation between r_t and β_t is given in equation 1.11 . At time t, we already have an estimate for β_{t-1} and a realization of the gaussian shock v_t. Having fixed matrix A we to determine β_t. Factor realizations at time t are put in H_t. From the gaussianity assumption for ϵ_t in the observation equation r_t is easily computed.

The a priori error estimates are :

$$e_t{}^* = \beta_t - \hat{\beta}_t^* \tag{1.15}$$

with covariance matrix :

$$P_t^* = E\left[e_t{}^* e_t^{*\prime}\right] \tag{1.16}$$

The posterior information set $I_t{}^*$ in addition to I_{t-1} contains the observation vector r_t and the a priori estimate $\hat{\beta}_t^*$:

$$I_t{}^* = \left\{r_t, r_{t-1}, ..., r_0, \hat{\beta}_t^*, \hat{\beta}_{t-1}, ..., \hat{\beta}_0\right\} \tag{1.17}$$

The a posteriori estimate for the state vector is obtained considering the a priori estimate and the return observation :

$$\hat{\beta}_t = E\left[\beta_t \,|I_t{}^*\right] \tag{1.18}$$

The a posteriori estimates for the estimation error and the covariance matrix are :

$$e_t = \beta_t - \hat{\beta}_t \tag{1.19}$$

$$P_t = E\left[e_t e_t'\right] \tag{1.20}$$

In the Kalman Filter system the a posteriori estimate $\hat{\beta}_t$ is supposed to be the a priori estimate

$\hat{\beta}_t^*$ plus a correction term which depends on K_t, the Kalman gain matrix :

$$\hat{\beta}_t = \hat{\beta}_t^* + K_t \left(r_t - H_t \hat{\beta}_t^* \right) \tag{1.21}$$

The Kalman gain matrix K_t is a measure of the modification from the a priori $\hat{\beta}_t$ to the a posteriori estimates after the observation of market returns r_t. If we are fully confident in our a priori estimates, that implies zero elements for P_t^*, the new observation r_t will not modify the estimation of the state space variables. Conversely, if the covariance matrix R of the measurement equation has all zero elements, our confidence in the measurements is high so the Kalman gain is high. Suppose now that we have already the a priori error and the corresponding covariance matrix. As soon as we get a new observation r_t, we update the estimation. After some simple calculations it is possible to have an expression for P_t in terms of the a priori error covariance matrix and the Kalman gain matrix:

$$\begin{aligned} P_t &= E\left[\left(\beta_t - \hat{\beta}_t \right) \left(\beta_t - \hat{\beta}_t \right)' \right] && (1.22) \\ &= E\left[\left(\beta_t - \hat{\beta}_t^* - K_t \left(r_t - H_t \hat{\beta}_t^* \right) \right) \left(\beta_t - \hat{\beta}_t^* - K_t \left(r_t - H_t \hat{\beta}_t^* \right) \right)' \right] && (1.23) \\ &= {P_t}^* - K_t H_t {P_t}^* - {P_t}^* H_t' K_t' + K_t H_t {P_t}^* H_t' K_t' + K_t R_t K_t' && (1.24) \end{aligned}$$

We used the fact that vectors e_t and ϵ_t are orthogonal. The intertemporal independence assumption for the errors implies that the error covariance matrix P_t is diagonal. The trace of P_t coincides with the sum of squared a posteriori errors. Observe that :

$$tr\left(P_t\right) = tr\left({P_t}^*\right) - tr\left(K_t H_t {P_t}^*\right) - tr\left({P_t}^* H_t' K_t'\right) + tr\left(K_t H_t {P_t}^* H_t' K_t'\right) + tr\left(K_t R_t K_t'\right) \tag{1.25}$$

Since P_t is the expectation of the squared residuals, we want to minimize it meaning that after the observation of return r_t the a posteriori error becomes small. The Kalman gain matrix introduced in equation 1.21 is chosen to be the solution of the following minimization problem:

$$\begin{cases} \min\limits_{K_t} tr(P_t) \\ s.t. \\ \quad \hat{\beta}_t = \hat{\beta}_t^* + K_t \left(r_t - H_t \hat{\beta}_t^* \right) \end{cases} , \tag{1.26}$$

Recall the following properties :

$$\frac{\partial tr(KX)}{\partial K} = X' \tag{1.27}$$

$$\frac{\partial tr(KBK')}{\partial K} = 2KB \tag{1.28}$$

where B is a symmetric matrix.

The minimum is obtained when :

$$\frac{\partial tr(P_t)}{\partial K} = 0 - (H_t P_t{}^*)' - P_t{}^* H_t' + 2K_t H_t P_t{}^* H_t' + 2K_t R_t \quad (1.29)$$

$$= 0 \quad (1.30)$$

The solution is the desired Kalman gain matrix :

$$K_t = P_t{}^* H_t' \left(H_t P_t{}^* H_t' + R_t\right)^{-1} \quad (1.31)$$

Using this matrix and the observation r_t, we are able to update the estimation:

$$\hat{\beta}_t^* = E\left[\beta_t \left| I_{t-1}\right.\right] = E\left[A\beta_{t-1} + v_{t-1} \left| I_{t-1}\right.\right] = A\hat{\beta}_{t-1} \quad (1.32)$$

It is now possible to compute the covariance error of the a priori estimator:

$$P_t^* = E\left[\left(\beta_t - \hat{\beta}_t^*\right)\left(\beta_t - \hat{\beta}_t^*\right)' | I_{t-1}\right] \quad (1.33)$$

$$= E\left[\left(A\left(\beta_{t-1} - \hat{\beta}_{t-1}\right) + v_{t-1}\right)\left(A\left(\beta_{t-1} - \hat{\beta}_{t-1}\right) + v_{t-1}\right)' | I_{t-1}\right] \quad (1.34)$$

$$= AP_{t-1}A' + Q \quad (1.35)$$

It is possible to have an explicit expression for P_t:

$$P_t = E\left[\left(\beta_t - \hat{\beta}_t\right)\left(\beta_t - \hat{\beta}_t\right)' | I_t^*\right] \quad (1.36)$$

$$= \left(I - K_t H_t\right) P_t{}^* \quad (1.37)$$

The Kalman filter algorithm gives an a priori estimate of the factor loading $\hat{\beta}_t^*$ and immediately after we observe the actual return it updates the estimate in order to improve its prediction in the next step. The two steps are called the forecast step where we get the a priori estimates for the state space variables and the corresponding error covariance matrix:

$$\hat{\beta}_t^* = A\hat{\beta}_{t-1} \quad (1.38)$$

$$P_t{}^* = AP_{t-1}A' + Q \quad (1.39)$$

and the update during which the a posteriori quantities are obtained :

$$K_t = P_t{}^* H_t' \left(H_t P_t{}^* H_t' + R_t\right)^{-1} \quad (1.40)$$

$$\hat{\beta}_t = \hat{\beta}_t^* + K_t\left(r_t - H_t\hat{\beta}_t^*\right) \quad (1.41)$$

$$P_t = \left(I - K_t H_t\right) P_t{}^* \quad (1.42)$$

From the description of the main passages needed for the algorithm construction it can be easily noticed that the initial conditions we put on two state space equations highly influence the solution and the difficulty in its attainment. In fact, the assumption of stability for the error terms ease the update of the covariance matrices for the a priori and the a posteriori estimate errors. However, the implementation of this algorithm is not always possible since, as we will see later, about 50 to 100 observations are needed in order to reach stability. For market risk factors this is not a problem while for macroeconomic factors where the data available are not so frequent we may be not able to implement the KF algorithm.

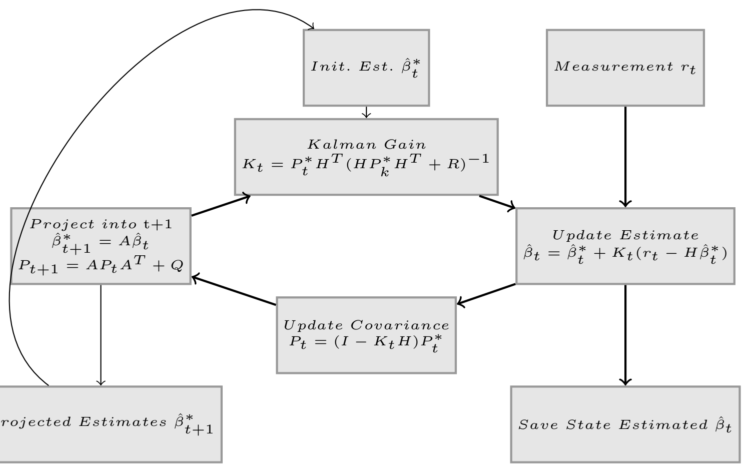


Figure 1.4: In this figure we describe how the Kalman Filter works. At time t, we have an a priori estimate for the factor exposure $\hat{\beta}_t^*$. With the available information we compute the Kalman Gain matrix K_t. The new observation r_t allows to update the state space variable estimation and get the a posteriori estimate $\hat{\beta}_t$. Now an a priori estimate for the state space variable at time $t+1$ becomes the starting point for the process to continue .

Multifactor models become dynamic by specifying the dynamic for each factor or by using time varying factor exposure β. The easiest method to obtain time - varying factor exposures is by using rolling Ordinary Least Squares regression. However, this approach is costly in terms of computations. It is preferable to use a more sophisticated instrument like Kalman Filter since it generates sequences for factor exposures that minimizes the variance of the prediction error as we discussed before.
We plot in figure 1.4 the OLS and KF factor loadings for the Vanguard Fund Utilities index Fund when three factors are considered: the Utilities Index, the exchange rate EURUSD and and the Industry Index. The OLS univariate regression gives a constant factor loading for the entire period, while the KF allows for time-varying exposures. In the first 100 iterations the algorithm is not so stable but after it gives reliable estimates. Observe that we supposed A=I in order to show that even the simplest choice can give time varying factor exposures. Q=8.5e-05 and R=7.9e-05 are maximum likelihood estimates based on conditional probabilities.

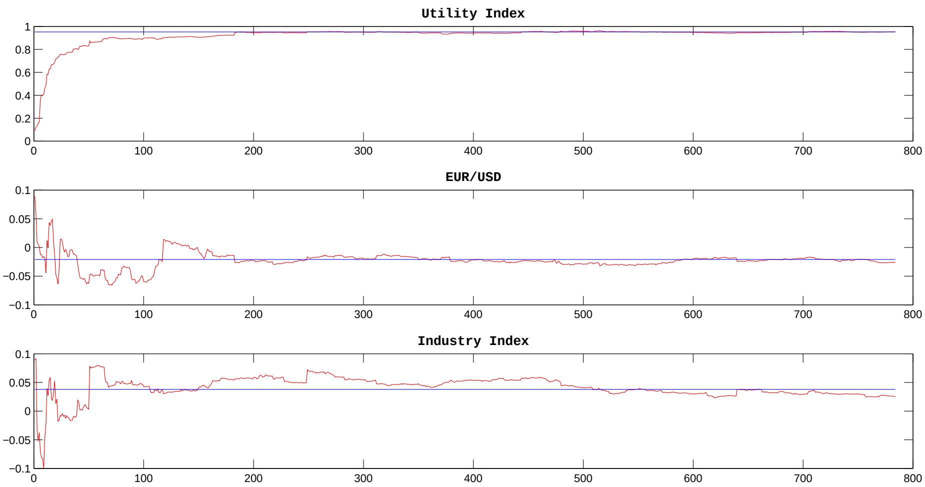


Figure 1.5: In the figure are shown results of univariate OLS regression and Kalman Filter analysis of the Vanguard Fund Utilities index vs three factors: Utility index, EURUSD and Industry Index.

Even if we use rolling OLS regression, the limit of having a mean factor exposure estimate for the entire period remains. On the other side, since the Kalman Filter algorithm is not so easy to understand its use in finance is limited.

1.2 Statistical Factors

Factor identification and dimensionality reduction are among the primary goals of multivariate financial time series analysis. They are useful in order to extract features and to find latent relations of risk drivers from high dimensional and complex portfolios. Attainment of these goals can be challenging.

Principal Component Analysis (PCA) is a mathematical procedure introduced in Pearson [1901] that uses an orthogonal transformation to convert a set of correlated variables into a set of uncorrelated variables called principal components (see Jolliffe [2002]). It yields uncorrelated factors but together with the Gaussian distribution assumption they become independent. In so doing, the problem of defining a dependence structure among the components is avoided.

Many papers test the performance of principal component analysis when applied to financial data . Different problems were faced with the ahead mentioned technique like hedging bond portfolios by Falkenstein and Hanweck [1997], interest rate forecasting by Reisman and Zohar [2004], immunization by Soto [2004] and value at risk estimation by Abad and Benito [2007]. PCA simultaneously solves factor identification and dimension reduction problems but the assumption of Gaussianity is not realistic for the description of financial data.

An alternative method for factor identification is Independent Component Analysis (ICA). ICA extracts Independent Components (ICs) but there is no projection on the maximum variance direction. This technique has been implemented in stock returns analysis by Back and Weigend [1997], in high frequency analysis by Moody and Wu [1998], in immunization problems by Bellini and Salinelli [2003],in an intertemporal GARCH context by Wu *et al.* [2006] and in risk management by Chen *et al.* [2010]. These articles show a nice performance of the ICA method, when applied to financial data that are not gaussianly distributed. However, the literature on ICA applied to finance is not so rich due to the fact that the factors do not have an economic interpretation.

The two methods, PCA and ICA, are different since they are based on quantities describing two dependence measures. Correlation is a measure of linear relation between variables while independence measures the existence of any relation between them. Note that independence, meaning that it doesn't exist any type of relation between variables, implies uncorrelation which considers only the linear relation between the variables. The vice versa is not always true.

Definition 1. *The covariance of two r.v's X and Y is :*

$$Cov(X,Y) = E[XY] - E[X]E[Y] \tag{1.43}$$

They are said to be uncorrelated if $Cov(X,Y) = 0$. The quantity :

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \tag{1.44}$$

is the correlation coefficient for the two r.v's X, Y. Notice that if the two r.v's are uncorrelated then :

$$E[XY] = E[X]E[Y] \tag{1.45}$$

Definition 2. *Two r.v's X, Y are said to be independent if their joint probability distribution is the product of their marginal probability distributions, i.e :*

$$P_{X,Y}(x,y) = P_X(x)P_Y(y) \tag{1.46}$$

for any realization x, y of the two r.v's.

From the Bayes rule we have that the conditional probability distribution for X is :

$$P_{X|Y}(x|y) = P_{X,Y}(x,y)/P_Y(y) \tag{1.47}$$

giving an equivalent definition for independence :

$$P_{X|Y}(x|y) = P_X(x) \tag{1.48}$$

Notice that since :

$$E[XY] = \int\int xyP_{X,Y}(x,y)dxdy = \int\int xyP_X(x)P_Y(y)dxdy \tag{1.49}$$

we have that independence implies uncorrelation :

$$E[XY] = \int xP_X(x)dx \int yP_Y(y)dy = E[X]E[Y] \tag{1.50}$$

The vice versa is not always true except for the case when X,Y are the marginals of a multivariate normal distribution.

1.2.1 Principal Component Analysis

There are many books explaining what Principal Component Analysis is. In particular Jolliffe [2002] gives in two sentences a general overview of the procedure :

The central idea of principal component analysis (PCA) is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. This is achieved by transforming to a new set of variables, the principal components (PCs), which are uncorrelated, and which are ordered so that the first few retain most of the variation present in all of the original variables.

PCA is used in multivariate analysis since it allows to obtain new uncorrelated variables. We describe the idea and some important characteristics of this technique. Consider a dataset given by a matrix $X_{p\times T}$ with p denoting the number of variables and T the length of time observation. In matrix notations each variable is given by a generic row :

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_i \\ \dots \\ x_p \end{bmatrix} = \begin{bmatrix} x_{1,1} & x_{1,2} & \dots & \dots & x_{1,T} \\ x_{2,1} & x_{2,2} & \dots & \dots & x_{2,T} \\ \dots & \dots & \dots & \dots & \dots. \\ x_{i,1} & x_{i,2} & \dots & \dots & x_{i,T} \\ \dots & \dots & \dots & \dots & \dots \\ x_{p,1} & x_{p,2} & \dots & \dots & x_{p,T} \end{bmatrix} \tag{1.51}$$

We want to generate a new set of uncorrelated variables by linearly transforming the original variables X. The new variables are the rows of the matrix $Y_{p\times T}$ obtained as :

$$Y = A'X \tag{1.52}$$

and are called Principal Components (PCs). Geometrically, they are the projections of the original variables on the rows of $A_{p\times p}$ that contain the eigenvectors associated to the covariance matrix Σ_X of the original variables defined as :

$$\Sigma_X = XX' \tag{1.53}$$

The covariance matrix Σ_X is symmetric and positive semidefinite[3]. Its eigenvalues are non negative and its eigenvectors are orthogonal. There exist infinitely many eigenvectors associated to each eigenvalue but by adding the unit length condition, the corresponding eigenvectors are orthonormal. The i-th principal component is obtained by multiplying the eigenvector associated to the i-th

[3] This is a prpoerty that satisfy all the covariance matrices

largest eigenvalue of the dataset covariance matrix with the original variables , i.e :

$$y_i = \begin{bmatrix} y_{i,1} & y_{i,2} & ... & y_{i,T} \end{bmatrix} = \begin{bmatrix} a_{1,i} & a_{2,i} & ... & a_{p,i} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_i \\ ... \\ x_p \end{bmatrix} \tag{1.54}$$

Putting together all the PCs we have :

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ ... \\ y_i \\ ... \\ y_p \end{bmatrix} = \begin{bmatrix} \sum_{i=1} a_{i,1}x_i \\ ... \\ \sum_{i=1} a_{i,k}x_i \\ ... \\ \sum_{i=1} a_{i,p}x_i \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{2,1} & ... & a_{i,1} & ... & a_{p,1} \\ a_{1,2} & a_{2,2} & ... & a_{i,2} & ... & a_{p,2} \\ ... & ... & ... & ... & ... & ... \\ a_{1,k} & a_{2,k} & ... & a_{i,k} & ... & a_{p,k} \\ ... & ... & ... & ... & ... & ... \\ a_{1,p} & a_{2,p} & ... & a_{i,p} & ... & a_{p,p} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_i \\ ... \\ x_p \end{bmatrix} \tag{1.55}$$

Uncorrelation for the transformed variables implies a diagonal form for the covariance matrix $\Sigma_Y = YY'$. The new and the original covariance matrices are connected since :

$$\Sigma_Y = A'X(A'X)' = A'XX'A = A'\Sigma_X A \tag{1.56}$$

Observe that Σ_X and Σ_Y are the covariance matrices respectively for the r.v's X and Y. The sample covariances are obtained by dividing the theoretical covariances with $T-1$.
The variance of the i-th PC is :

$$Var(y_i) = a_i'\Sigma_X a_i \tag{1.57}$$

and it coincides with λ_i that is the i-th largest eigenvalue of the matrix Σ_X. The uncorrelation of the PCs allows to linearly decompose the total variance, i.e :

$$Var(Y) = Var(y_1) + Var(y_2) + ... + Var(y_p) = \lambda_1 + \lambda_2 + ... + \lambda_p. \tag{1.58}$$

The PCA is known as a method used to reduce the number of variables considered in order to make easier the identification of common features among the original variables. This is not done without an information loss but in some circumstances with only few PCs we can reproduce a large portion of the original variability. Considering only the first r PCs ordered according to their variance, with r < p, the explained variance is :

$$\frac{\lambda_1 + \lambda_2 + ... + \lambda_r}{\lambda_1 + \lambda_2 + ... + \lambda_p} \tag{1.59}$$

We can go on two directions: fix r (the number of PCs) and compute the variance explained by the first r PCs or choose first the level of the variance to be explained and then find the minimum number of PCs for which it is possible. In a financial context, we aim at explaining most of the variance present in the dataset so we follow the second direction. Eigenvectors are calculated iteratively. The two most applied procedures are the singular value decomposition (SVD) and Jacobi Rotation method. The last one is slower but more accurate. See Golub and Van Loan [1996] for more details on these procedures.
PCs are uncorrelated variables each orthogonal to the previous one. Let us see how the PCA works for a given dataset X:

$$X = (x_1, x_2, ..., x_n) \tag{1.60}$$

This procedure can be represented as a constrained maximization problem. Each PC is a convex combination of the observations:

$$y_k = \sum_{i=1}^{p} a_{ik} x_i = a_k^{'} X \tag{1.61}$$

where a_k, for any $k > l \geq 1$ is the solution of the problem:

$$\begin{cases} \max\limits_{a_k} Var\left(y_k\right) \\ s.t. \\ Cov\left(y_k, y_l\right) = 0 \\ a_k^{'} a_k = 1 \end{cases}, \tag{1.62}$$

The consequence of zero covariance is the orthogonality between the PCs.
Let us see more details on how the PCs are computed. In particular, to find a_1 we must observe that:

$$Var\left(y_1\right) \quad = \quad Var\left(x_1 a_{11} + x_2 a_{21} + ... + x_n a_{n1}\right) = a_1^{'} Cov\left(x_1, ..., x_n\right) a_1 = a_1^{'} \Sigma a_1 \tag{1.63}$$

so that the maximization problem becomes:

$$\begin{cases} \max\limits_{a_1} Var\left(y_1\right) \\ s.t. \\ a_1^{'} a_1 = 1 \end{cases}, \tag{1.64}$$

By using the technique of Lagrange multipliers we can rewrite the problem in the form :

$$\underset{a_1}{max} \left(a_1^{'} \Sigma a_1 - \lambda_1 \left(a_1^{'} a_1 - 1\right)\right) \Leftrightarrow \Sigma a_1 - \lambda_1 a_1 = 0 \tag{1.65}$$

We recognize that this problem can be viewed as an eigenvalue decomposition since we have:

$$\Sigma a_1 = \lambda_1 a_1 \tag{1.66}$$

Notice that a_1 is an eigenvector of Σ associated to the eigenvalue λ_1. Furthermore, this eigenvalue coincides with the variance of the first component. Indeed, the variance associated to the vector y_1 is:

$$Var\left(y_1\right) = a_1^{'}\Sigma a_1 = a_1^{'}\lambda_1 a_1 = \lambda_1 \tag{1.67}$$

Eigenvector a_2 is the solution of the following constrained problem:

$$\begin{cases} \max\limits_{a_2} Var\left(y_2\right) \\ s.t. \\ Cov\left(y_2, y_1\right) = 0 \\ a_2^{'}a_2 = 1 \end{cases}, \tag{1.68}$$

The uncorrelation between the first two PCs gives a condition about the product of the eigenvectors a_1 and a_2:

$$Cov\left(a_1^{'}X, a_2X\right) = a_1^{'}\Sigma a_2 = a_2^{'}\Sigma a_1 = a_2^{'}\underbrace{\lambda_1 a_1}_{\Sigma a_1} = \lambda_1 a_2^{'}a_1 = 0 \Leftrightarrow a_2^{'}a_1 = 0 \tag{1.69}$$

By writing the problem in the Lagrangian form we have:

$$\underset{a_2}{max}\left(a_2^{'}\Sigma a_2 - \lambda_2\left(a_2^{'}a_2\right) - \phi\left(a_2^{'}a_1\right)\right) \Leftrightarrow \Sigma a_2 - \lambda_2 a_2 - \phi a_1 = 0 \Leftrightarrow a_1^{'}\Sigma a_2 - \lambda_2 a_1^{'}a_2 - \phi a_1^{'}a_1 = 0 \tag{1.70}$$

The first two quantities are zero, so it remains that:

$$-\phi\underbrace{a_1^{'}a_1}_{1} = 0 \Leftrightarrow \phi = 0 \tag{1.71}$$

The solution of the Lagrangian problem can be found in the eigenvalue problem:

$$\Sigma a_2 - \lambda_2 a_2 = 0 \Leftrightarrow \Sigma a_2 = \lambda_2 a_2 \tag{1.72}$$

In the same way, it can be shown that the maximum variance for the i-th PC score is i-th largest eigenvalue.

In real world applications we face the problem of deciding whether the Principal Component Analysis should be applied on the covariance or on the correlation matrix of the given dataset. Results are affected by the volatility of each variable when you do the PCA on the covariance matrix, whereas they are independent of volatility when PCA is based on the correlation matrix. Applying PCA to covariance or correlation is an empirical question.
Note that PCA is not scale invariant and using the correlation matrix for PCA means to rescale the data such that they all have unit variance. This feature suggests that results of PCA on covariance and correlation matrices could differ significantly. If assets in the portfolio are of the

same type and with similar variance, the choice is not so relevant. But if we consider a portfolio composed of bonds and equities with bond yields being much less volatile than equity returns, PCA on covariance matrix would give higher weights to equities in the first component than if we use the correlation matrix. Significant differences in variable variances implies a greater influence of the variables with the largest variance in the first eigenvectors associated to the covariance matrix. However, sometimes using PCA on correlation matrix may be convenient. For example, consider a portfolio composed by a bond and a stock where bond return is expressed in basis points and stock return is expressed in units. Applying directly PCA on the covariance matrix would give misleading results since the two variables are not measured using the same unit. In some sense choosing to work with correlation matrices protects from these kind of problems but has the drawback that not all the information present in the data is used.

We will follow a hybrid approach composed by two steps. In the first step we standardize the variables by using the volatility of a previous fixed period that can be for example the volatility of the last quarter. Depending on the past, the relative variances are near one but they can be higher or lower since the market conditions change frequently. As a second step we perform PCA on the new covariance matrix whose diagonal elements will have similar values. In doing so, we maintain some of the information about greater volatility of some assets and in the same time have no big differences in the data units.

In particular we consider the five GICS indexes: Information Technology, Financial, Consumption Staple, Health and Energy. The dataset is composed by daily log returns going from 12/09/2010 to 12/09/2012. We first perform a PCA analysis on the covariance matrix computed using data from 01/01/2011 to 12/09/2012.

For the same time frame we perform the PCA on the correlation matrix. It is possible to observe that now the variance explained from the first PC is lower than in the case where the covariance matrix. We compare the results with the hybrid approach. The volatility computed on the period 12/09/2010 - 31/12/2012 is used to standardize the data in the remaining period. The outcome is coherent with our expectations, in the sense that the first PC explains more than in the correlation case and less than in the covariance context.

Cov	Inf	Fin	Cons	Hlth	Enr	Eigenvalues	Perc
Inf	**1.56E-04**	1.74E-04	7.16E-05	9.79E-05	1.53E-04	7.05E-04	**86%**
Fin	1.74E-04	**2.91E-04**	9.72E-05	1.32E-04	2.11E-04	4.69E-05	92%
Cons	7.16E-05	9.72E-05	**5.62E-05**	6.14E-05	8.66E-05	3.44E-05	96%
Hlth	9.79E-05	1.32E-04	6.14E-05	**9.21E-05**	1.17E-04	2.50E-05	99%
Enr	1.53E-04	2.11E-04	8.66E-05	1.17E-04	**2.27E-04**	1.01E-05	100%

Table 1.1: Covariance matrix using five sector indexes going from January 2011 to September 2012. The Consumer Staple and Health sectors show a lower variance meaning that the first eigenvalues will be driven by the remaining sectors are Information Technology, Financial and Energy.

Corr	Inf	Fin	Cons	Hlth	Enr	Eigenvalues	Perc
Inf	**1.00**	0.82	0.77	0.82	0.82	4.22	**84%**
Fin	0.82	**1.00**	0.76	0.81	0.82	0.28	90%
Cons	0.77	0.76	**1.00**	0.85	0.77	0.19	94%
Hlth	0.82	0.81	0.85	**1.00**	0.81	0.18	97%
Enr	0.82	0.82	0.77	0.81	**1.00**	0.14	100%

Table 1.2: Correlation matrix using the data going from January 2011 to September 2012. The explained variance by the first PC is lower than in the case when covariance matrix was used.

Hybrid	Inf	Fin	Cons	Hlth	Enr	Eigenvalues	Perc
Inf	**1.02**	0.82	0.75	0.88	0.88	4.42	**85%**
Fin	0.82	**1.00**	0.75	0.87	0.88	0.28	90%
Cons	0.75	0.75	**0.94**	0.89	0.79	0.20	94%
Hlth	0.88	0.87	0.89	**1.10**	0.91	0.19	98%
Enr	0.88	0.88	0.79	0.91	**1.14**	0.11	100%

Table 1.3: Covariance matrix using hybrid data : the last quarter of 2010 was used to compute the basis standard deviation . The data going from January 2011 to September 2012 were then "standardized" and used for the covariance matrix computation. The explained variance by the first PC is lower than when we use covariance matrix for PCA and higher than in the correlation matrix case.

Now we apply the PCA analysis on Euro market Z-spread data grouped by sectors and subsectors in order to identify possible leading factors. The notion of spread is encountered frequently in finance, especially in fixed income. We call nominal spread the difference between the yield to maturity of a risky bond and the yield to maturity of a risk-free one with the same maturity.

The zero volatility spread (z-spread) is the fixed amount that should be added to the Libor curve in order to obtain the market price for a given bond. In the last decade credit risk has become very important. In fact for portfolios that contain bonds, the z-spreads are crucial risk factors. If we are in the market and have in the portfolio positions on a large number of bonds, identifying a smaller number of factors that can explain the movements of the z-spreads for different ratings reduces the number of risk factors considered in the portfolio management. In addition, the information we get can be used to identify the sectors or subsectors that have a higher influence for a given rating.

In the considered market we have that most of the investment grade issuers are AA, A or BBB rated. As uncertainty grows, in the market there is much more reluctance in assigning an AAA rating to a given issuer. Even for the A rated is not easy to have a large dataset since a lot of companies originally having this rating have been either uprated or more probably downrated. As we will show for the A group the most representative sector is the Public Sovereign since it is perceived as safer than other sectors whose issuers are private companies. However we were able to perform a complete analysis only for the time span going from 01/08/2011 to 29/02/2012. The considered period is not too long but it is interesting to observe the most important factors driving the z-spread of the *AA* rated companies.

	Util Gas	Util Transp	Fin Bank	Util Water	Fin Finance	Fin Insurance	Basic Industries	Indus Healthcare	Indus Manifact	Indus Retail	Indus Technology	Indus Telecom	Pub Sovereign	Util Electric
Util_Gas	**1.410**	1.209	0.795	1.343	0.865	0.623	0.818	0.944	0.853	0.963	0.546	0.936	0.302	1.351
Util_Transp	1.209	**1.037**	0.681	1.152	0.742	0.534	0.702	0.809	0.732	0.826	0.468	0.803	0.259	1.159
Fin_Bank	0.795	0.681	**1.648**	0.753	1.610	0.757	0.374	0.432	0.430	0.497	0.181	0.678	0.035	0.761
Util_Water	1.343	1.152	0.753	**1.291**	0.823	0.592	0.780	0.902	0.812	0.917	0.519	0.895	0.299	1.287
Fin_Finance	0.865	0.742	1.610	0.823	**2.466**	0.886	0.375	0.474	0.457	0.507	0.157	0.681	0.198	0.829
Fin_ Insurance	0.623	0.534	0.757	0.592	0.886	**0.883**	0.351	0.387	0.375	0.371	0.263	0.403	0.266	0.597
Basic_Industries	0.818	0.702	0.374	0.780	0.375	0.351	**0.937**	0.720	0.675	0.765	0.518	0.680	0.233	0.784
Indus_Healthcare	0.944	0.809	0.432	0.902	0.474	0.387	0.720	**1.007**	0.765	0.892	0.517	0.706	0.230	0.904
Indus_Manufacturing	0.853	0.732	0.430	0.812	0.457	0.375	0.675	0.765	**0.909**	0.824	0.490	0.669	0.160	0.817
Indus_Retail	0.963	0.826	0.497	0.917	0.507	0.371	0.765	0.892	0.824	**1.239**	0.499	0.807	0.229	0.922
Indus_Technology	0.546	0.468	0.181	0.519	0.157	0.263	0.518	0.517	0.490	0.499	**1.275**	0.358	0.124	0.523
Indus_Telecom	0.936	0.803	0.678	0.895	0.681	0.403	0.680	0.706	0.669	0.807	0.358	**1.101**	0.158	0.897
Pub_Sovereign	0.302	0.259	0.035	0.299	0.198	0.266	0.233	0.230	0.160	0.229	0.124	0.158	**2.953**	0.290
Util_Electric	1.351	1.159	0.761	1.287	0.829	0.597	0.784	0.904	0.817	0.922	0.523	0.897	0.290	**1.294**

Table 1.4: Rated A subsectors : Covariance matrix using the 2010 historical standard deviation as basis for standardizing the data.

The corresponding eigenvalues and explained variance are given in the table 1.10:

By observing the eigenvectors we can find the sectors that drive the market.

If we choose to build the covariance matrix based on the hybrid data with base the entire year we observe still big differences in terms of variability among the subsectors. A more accurate

Eigenvalues	Explained Variance
10.293	**53%**
2.955	**68%**
2.717	**82%**
0.979	**87%**
0.668	**91%**
0.430	93%
0.423	95%
0.321	97%
0.257	98%
0.214	99%
0.183	100%
0.009	100%
0.000	100%
0.000	100%

Table 1.5: Eigenvalues associated to the A rated covariance matrix using 2010 as basis year

	Eigvec1	Eigvec2	Eigvec3	Eigvec4	Eigvec5
Util_Gas	0.353	0.028	0.101	0.163	0.317
Util_Transp	0.302	0.024	0.087	0.140	0.271
Fin_Bank	0.263	**-0.325**	**-0.369**	-0.066	-0.075
Util_Water	0.336	0.031	0.096	0.159	0.304
Fin_Finance	0.302	**-0.386**	**-0.572**	-0.144	-0.174
Fin_ Insurance	0.191	-0.079	-0.190	-0.143	0.236
Basic_Industries	0.230	0.094	0.174	-0.077	-0.248
Indus_Healthcare	0.259	0.083	0.180	-0.001	-0.208
Indus_Manufacturing	0.239	0.055	0.168	-0.036	-0.261
Indus_Retail	0.275	0.081	0.193	0.048	**-0.520**
Indus_Technology	0.160	0.104	0.218	**-0.908**	0.153
Indus_Telecom	0.265	-0.013	0.079	0.160	-0.287
Pub_Sovereign	0.106	**0.837**	**-0.531**	-0.005	-0.034
Util_Electric	0.338	0.027	0.097	0.156	0.303

Table 1.6: Eigenvectors of the A rated sectors

procedure could be based on considering only the last quarter of the year 2010 since volatility rose to higher levels in all the sectors. The standard deviation computed on the entire year gives a flatter measure than using only the more volatile period. We decided to standardize the 2011-2012 dataset using the volatile period and compare the results with the A rated group :

	Util Gas	Util Transp	Fin Bank	Util Water	Fin Finance	Fin Insurance	Basic Industries	Indus Healthcare	Indus Manifact	Indus Retail	Indus Technology	Indus Telecom	Pub Sovereign	Util Electric
Util_Gas	**1.096**	0.984	0.648	1.121	0.763	0.582	0.805	0.814	0.789	0.832	0.405	1.102	0.234	1.127
Util_Transp	0.984	**0.883**	0.582	1.006	0.685	0.523	0.723	0.730	0.709	0.747	0.364	0.989	0.210	1.012
Fin_Bank	0.648	0.582	**1.411**	0.659	1.490	0.743	0.387	0.391	0.417	0.450	0.141	0.837	0.028	0.666
Util_Water	1.121	1.006	0.659	**1.156**	0.779	0.594	0.824	0.834	0.806	0.851	0.414	1.130	0.248	1.152
Fin_Finance	0.763	0.685	1.490	0.779	**2.468**	0.940	0.419	0.464	0.480	0.497	0.132	0.909	0.174	0.784
Fin_ Insurance	0.582	0.523	0.743	0.594	0.940	**0.993**	0.416	0.402	0.417	0.386	0.235	0.571	0.247	0.599
Basic_Industries	0.805	0.723	0.387	0.824	0.419	0.416	**1.168**	0.786	0.791	0.837	0.487	1.013	0.228	0.828
Indus_Healthcare	0.814	0.730	0.391	0.834	0.464	0.402	0.786	**0.963**	0.785	0.855	0.426	0.921	0.197	0.837
Indus_Manufacturing	0.789	0.709	0.417	0.806	0.480	0.417	0.791	0.785	**1.001**	0.847	0.433	0.937	0.147	0.812
Indus_Retail	0.832	0.747	0.450	0.851	0.497	0.386	0.837	0.855	0.847	**1.189**	0.411	1.055	0.197	0.855
Indus_Technology	0.405	0.364	0.141	0.414	0.132	0.235	0.487	0.426	0.433	0.411	**0.904**	0.403	0.091	0.417
Indus_Telecom	1.102	0.989	0.837	1.130	0.909	0.571	1.013	0.921	0.937	1.055	0.403	**1.962**	0.185	1.133
Pub_Sovereign	0.234	0.210	0.028	0.248	0.174	0.247	0.228	0.197	0.147	0.197	0.091	0.185	**2.272**	0.240
Util_Electric	1.127	1.012	0.666	1.152	0.784	0.599	0.828	0.837	0.812	0.855	0.417	1.133	0.240	**1.159**

Table 1.7: Rated A subsectors : Covariance matrix using the last quarter of 2010 standard deviation.

Eigenvalues	Explained Variance
10.193	**55%**
2.700	**69%**
2.223	**81%**
0.807	**85%**
0.690	**89%**
0.513	**92%**
0.447	94%
0.340	96%
0.297	98%
0.226	99%
0.180	100%
0.008	100%
0.000	100%
0.000	100%

Table 1.8: Eigenvalues associated to the A rated covariance matrix using the last quarter of 2010 for the standard deviation used for the hybrid data

Now to explain at least 90 per cent of the total variability we need six factors, one more than the factors needed when we use the entire year 2010 for the estimation. The aim of PCA is to reduce the number of factors through which we can reproduce most of the variability originally observed in the market. At a first glance it seems that we are going on the wrong direction. If we concentrate only on the number of PCs yes, but if we look behind the second method gives more information. Indeed, by looking at the second and third eigenvectors we observe that now each PC, excluding the first one which gives almost equal weight, is dominated by one specific sector.

The second PC explains 14 per cent of the total variability and gives more weight to the finance sector. In particular it seems strongly influenced by the Financial Bank and Financial Finance subsectors. The third PC explains 12 per cent of the total variability and is highly dependent on the Public Sovereign subsector. The next two PCs explain each 4 per cent of the total variability but are driven by different sectors. The fourth PC influenced mostly by the Industrial sectors gives a quite big weight to the Industrial Technology subsector and negative weight to the Industrial Telecomunication. The fifth PC gives negative weight to the Industrial and Financial sectors while the dependence from the Utility seems strongly positive. Changing the time frame used for the estimation of the standard deviation we obtain different results.

	Eigvec1	Eigvec2	Eigvec3	Eigvec4	Eigvec5	Eigvec6
Util_Gas	0.310	0.079	0.027	0.006	0.357	0.037
Util_Transp	0.278	0.071	0.024	0.005	0.320	0.033
Fin_Bank	0.244	**-0.459**	0.004	-0.034	**-0.094**	0.179
Util_Water	0.317	0.083	0.024	0.001	0.367	0.040
Fin_Finance	0.302	**-0.706**	-0.098	0.066	**-0.197**	**-0.285**
Fin_ Insurance	0.202	-0.221	-0.094	0.320	0.172	0.187
Basic_Industries	0.262	0.224	0.042	0.093	**-0.280**	**-0.057**
Indus_Healthcare	0.255	0.187	0.046	0.079	**-0.067**	**-0.304**
Indus_Manufacturing	0.255	0.173	0.070	0.090	**-0.163**	**-0.299**
Indus_Retail	0.273	0.200	0.062	-0.063	**-0.272**	**-0.450**
Indus_Technology	0.132	0.173	0.041	**0.708**	**-0.336**	0.432
Indus_Telecom	0.368	0.085	0.098	**-0.600**	**-0.354**	0.521
Pub_Sovereign	0.085	0.158	**-0.977**	-0.058	**-0.053**	0.019
Util_Electric	0.319	0.081	0.028	0.006	0.367	0.038

Table 1.9: Eigenvectors of the A rated sectors obtained with the hybrid data whose base standard deviation is computed using the last quarter of 2010

We perform the same analysis for the BBB rated issuers. We use the hybrid approach based on the standard deviation estimated during the last quarter of 2010. By looking at the eigenvectors, we observe that now we cannot claim that the PCs give a higher weight to a specific sector i.e each new factor is now a weighted average of the different sectors. We now need eight factors to explain 90 per cent of the total variability. The variance explained by each PC is not very large . Unfortunately there is not a unique recipe that tells us how long we must go back in the estimation period and how long these estimates can be used. In turbulent periods, volatility changes rapidly so the covariances of the hybrid data will probably be quite different from one. On the other side if the market is quiet volatility moves slowly so the covariances obtained from the hybrid data will be near one. If we observe a severe change of the market conditions probably this is due to one or more factors in particular so the update of the basis standard deviation becomes compulsory. However, even if market conditions remain stable, the factors driving the market may change. We

Eigenvalues	Explained Variance
16.721	**48%**
3.908	**59%**
3.215	**68%**
2.507	**75%**
1.696	**80%**
1.521	**84%**
1.259	**88%**
0.955	**90%**
0.850	93%
0.736	95%
0.594	97%
0.464	98%
0.417	99%
0.285	100%
0.022	100%
0.000	100%
0.000	100%

Table 1.10: Eigenvalues associated to the BBB rated covariance matrix using the last quarter of 2010 for the standard deviation estimation

	Eigvec1	Eigvec2	Eigvec3	Eigvec4	Eigvec5	Eigvec6	Eigvec7	Eigvec8
Utility Gas_Pipelines	0.310	0.233	0.124	0.141	0.028	0.016	0.037	0.031
Utility Reg_Transp	0.445	0.334	0.178	0.203	0.041	0.023	0.054	0.045
Utility Water	0.298	0.225	0.118	0.130	0.031	0.001	0.040	0.030
Utility Electric	0.300	0.225	0.120	0.136	0.027	0.015	0.036	0.030
Financial Bank	0.182	0.004	0.126	-0.193	**-0.163**	0.093	0.087	0.095
Financial Finance	0.252	**-0.175**	0.177	**-0.364**	**-0.664**	**0.367**	-0.045	-0.091
Financial Insurance	0.340	**-0.272**	0.159	**-0.620**	**0.354**	**-0.470**	0.137	-0.012
Industrial Basic_Industries	0.168	**-0.167**	0.004	-0.004	-0.037	0.255	-0.014	-0.054
Industrial Consumer_Products	0.147	**-0.244**	-0.159	0.166	0.133	0.107	-0.052	-0.032
Industrial Energy	0.232	0.075	0.030	0.116	0.063	-0.189	-0.317	-0.415
Industrial Manufacturing	0.248	**-0.235**	-0.063	0.042	0.053	-0.010	-0.233	0.251
Industrial Media	0.125	**-0.250**	-0.138	0.138	0.131	0.061	-0.004	-0.290
Industrial Property	0.142	**-0.268**	-0.205	0.368	**-0.531**	**-0.595**	0.268	-0.015
Industrial Retail	0.234	**-0.379**	-0.193	0.180	0.137	0.176	-0.335	**0.460**
Industrial Technology	0.109	**-0.204**	-0.108	0.076	0.137	0.236	0.069	**-0.651**
Industrial Telecom	0.194	0.387	**-0.832**	**-0.330**	-0.076	0.004	-0.057	-0.006
Industrial Transportation	0.097	-0.103	-0.141	0.067	0.183	0.280	**0.784**	0.140

Table 1.11: The first 8 eigenvectors associated to the covariance matrix of the BBB rated issuers

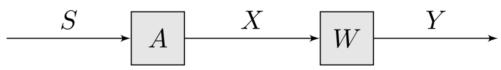
face the same problem as in curve construction or market timing problems: it is not a science but a matter of art and making the right decision is left to the sensitivity of the subject based on its past experience.

1.2.2 Independent Component Analysis

Independent component analysis (ICA) (see Hyvarinen [1999]) is a mathematical procedure for revealing hidden factors driving the realizations of the observed variables. As in the PCA case, we have the multivariate data matrix $X_{p\times T}$ and we want to find the independent components that are the rows of matrix $S_{p\times T}$. Again we suppose that the observed variables are a linear transformation of the original factors through a matrix $A_{p\times p}$, i.e :

$$X = AS \tag{1.73}$$

Except for the linearity all the other assumptions presented in PCA are relaxed. In this model the original sources S, called independent components (ICs), are mixed with matrix A giving the observations in X.



In fact, we do not find directly matrix A and the unobservable variables in S but we look for a demixing matrix $W_{p\times p}$ from which we obtain $Y_{p\times T}$ whose rows are the estimates for the unobservable sources.

$$Y = WX \tag{1.74}$$

Matrix W is obtained using any ICA algorithm which performs an optimization procedure whose objective function depends on a measure of independence of the ICs. Variance maximization is now substituted with independence for finding the hidden factors.

In literature we can find different approaches and related algorithms for the application of this model. ICA can be derived in various ways, for example by maximization of non gaussianity, maximum likelihood estimation or minimization of mutual information. Although there are some recent papers on parametric approaches (see for example Lu *et al.* [2009] for an example on parametric ICA for modeling financial time series), most of the literature is about non-parametric ICA.

In the general definition, each vector Y_i represents an independent component that is actually an estimate of the i-th source S_i. There is an ambiguity in the ICA method in the sense that an estimated signal Y_i determines up to a multiple the variance of the source S_i. That is, there exist infinite values for α_i such that this decomposition holds :

$$X_i = \tilde{A}_i \tilde{S}_i \tag{1.75}$$

where $\tilde{A}_i$ is obtained by dividing with α_i the corresponding elements in matrix A while we must multiply the elements in S with α_i to obtain the elements in $\tilde{S}_i$. It is possible to choose α_i in that way that we have a unit variance signal but the sign ambiguity still remains.

Another problem is the fact that it is not possible to uniquely determine the order of the independent components. The reason is that given a permutation matrix P and its inverse the model can be decomposed in:

$$X = AP^{-1}PS \tag{1.76}$$

Any algorithm chosen for factor identification can suppose that the original independent variables are PS instead of S with estimated mixing matrix being AP^{-1}.
It is reasonable to affirm that the observed variables are more gaussian than the distribution of some factors. Indeed, if we think at the observed variables as a linear transformation of more than two r.v's, whose distribution is unknown for us, from the Central Limit Theorem we have that the univariate distributions can be highly non gaussian.

As we already mentioned one way of estimating the ICs we can maximize a measure of nonguassianity for each component. FastICA is an algorithm (see Hyvarinen [1999]) based on a fixed-point iteration scheme maximizing non-Gaussianity. Since we will use it in many application let us recall the main passages in its construction.

The ICs are assumed to be non-gaussian. A measure of non-gaussianity is the negentropy $J(Y)$ defined as a difference of two quantities :

$$J(y) = H(\hat{y}) - H(y) \tag{1.77}$$

where $\hat{y}$ is a random vector coming from a gaussian distribution while the random vector y can be generated using any distribution. Observe that $H(y)$ is a measure of entropy for the random vector y :

$$H(y) = \int -p(y) log p(y) dy \tag{1.78}$$

The negentropy can be seen as a measure of the distance from the perfect disorder of the gaussian world. It is possible to use an approximation for the negentropy which facilitates computations:

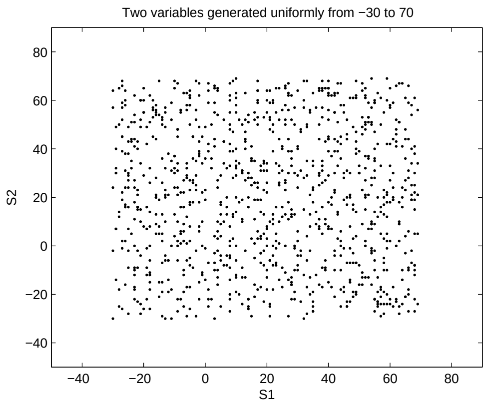
$$J(y) \approx \left[E\left\{G(\hat{y})\right\} - E\left\{G(y)\right\}\right]^2 \tag{1.79}$$

for particular choices of function G. In particular, we consider choose :

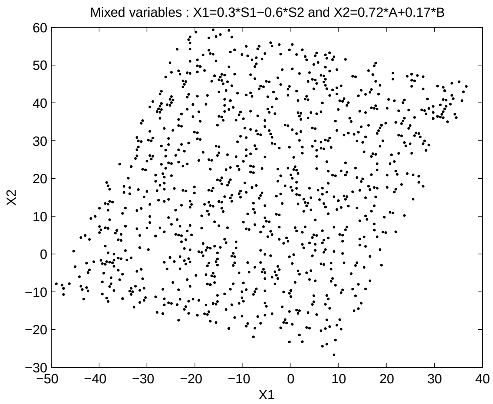
$$G(Y) = e^{-y^2/2} \tag{1.80}$$

Before applying the ICA algorithm the original dataset must be preprocessed, i.e centered and whitened. Consider a dataset composed by two r.v's uniformly distributed on $[-30, 70]$. Each original variable is centered meaning that we subtract the vector mean to the dataset. The centered data are then whitened, i.e the transformed dataset $\hat{X}$ is such that:

$$E[\hat{X}\hat{X}'] = I \tag{1.81}$$



(a) True signals



(b) Observed variables

Figure 1.6: In the left figure we plot the two original signals $S1$ and $S2$ that are uniformly distributed on $[-30, 70]$. The observed quantities $X1$ and $X2$ are linear transformations of the signals S1 and S2. The plot in the right refers to $X1 = 0.3 * S1 - 0.6 * S2$ and $X2 = 0.72 * S1 + 0.17 * S2$.

A common way to trasform the data is by using the matrix obtained from the eigenvalue and eigenvectors decomposition of $Cov(X) = E[XX']$. If we put the eigenvectors in matrix E and D is a diagonal matrix whose elements are the corresponding eigenvalues $D = diag(\lambda_1, \lambda_2, .., \lambda_n)$ then

$$X = EDE' \tag{1.82}$$

The performed transformation :

$$\hat{X} = ED^{-\frac{1}{2}}E'X \tag{1.83}$$

yields the whitened data. Since :

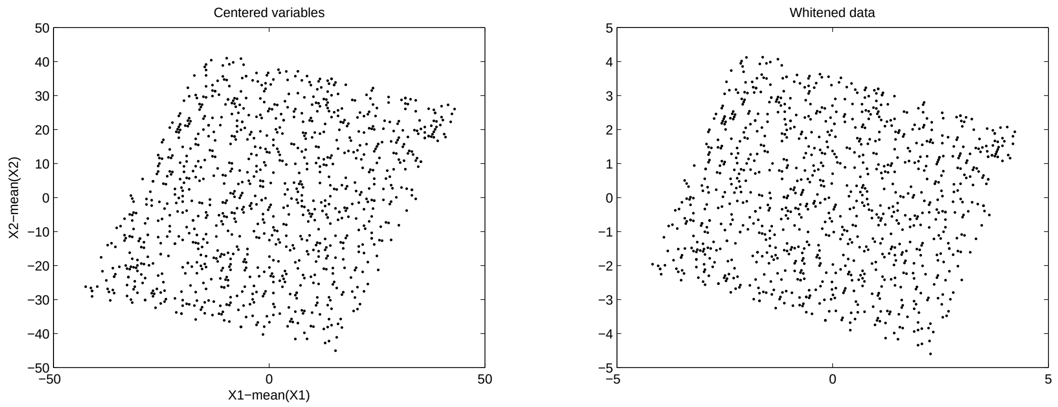
$$\hat{X} = ED^{-\frac{1}{2}}E'AS = \hat{A}S \tag{1.84}$$

the new mixing matrix $\hat{A}$ is orthogonal because :

$$E[\hat{X}\hat{X}'] = \hat{A}E[SS']\hat{A}' = \hat{A}\hat{A}' = I \tag{1.85}$$

FastICA is an iterative algorithm and it finds the direction for the weight vector w maximizing the non-Gaussianity of the projection w' for the data $\hat{x}$. The function $g(\cdot)$ is the derivative of the function G defined in equation 1.80. The main steps of the algorithm to obtain one IC are:

- 1 Start with a random weight vector w
- 2 Define $w^+ \leftarrow E\left\{\hat{x}g(w^{'}\hat{x})\right\} - E\left\{g'(w^{'}\hat{x})\right\} w$



(a) Centered variables (b) Observed variables

Figure 1.7: Preprocessing the dataset means to center the data and then whiten them. The last term means that the covariance matrix of the transformed dataset is diagonal.

- 3 Divide $w \leftarrow w^+/\|w^+\|$
- 4 Compare w with the previous one. If the difference is bigger than a fixed value ϵ, it means that there is no convergence so the update the value as in 2.

In multivariate analysis the identification of the main factors that generate the variables that we observe is an important step. Under the stationary hypothesis it gives the possibility to use the past information to predict the future. Here we apply the ICA analysis and try to identify the factors driving the returns on stock markets. We consider the ten GICS sectors as factors for the VFIAX index and apply the FastICA algorithm to find 10 independent components . In 1.2.2 differences between the components can be immediately noticed.

Madan [2006a] suggests to consider as independent factors only those that differ substantially from the normal distribution while the remaining components are grouped together to form the idiosyncratic component. The difficulty in the component interpretation explains the limited use of this instrument in finance although from a mathematical point of view it is very interesting.

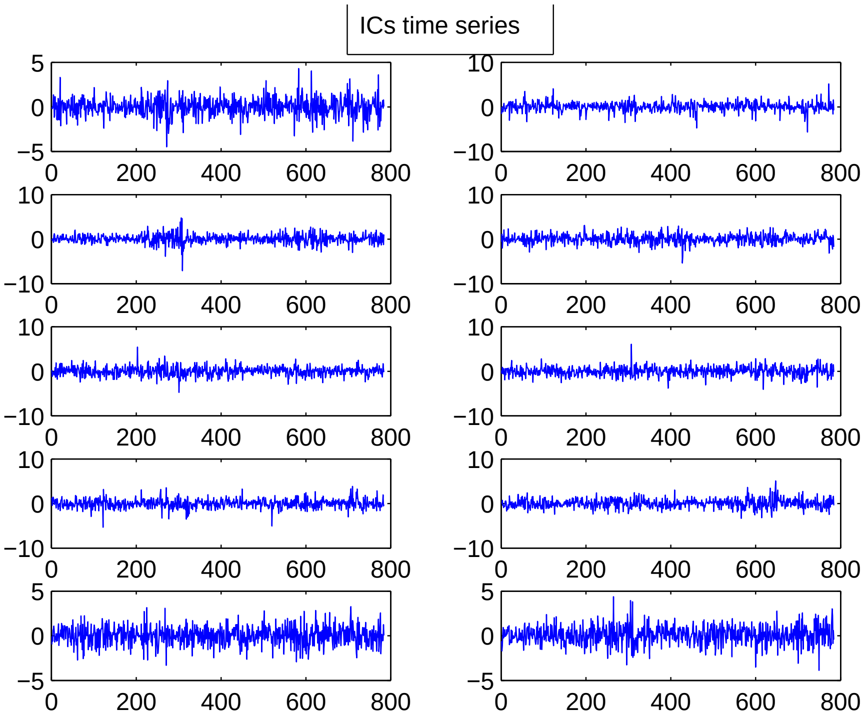


Figure 1.8: We consider the ten GICS sectors as factors for the VFIAX index and apply the FastICA algorithm to find 10 ICs that generate our portfolio return. In the figure we show the time series of the ICs

1.2.3 PCA vs ICA

PCA and ICA yield both linear transformations of the variables but they are based on quite different hypothesis and yield components that have different statistical properties. In fact we have that PCs are uncorrelated and gaussian while ICs are independent and supposed non gaussian . Only Second-order statistics are needed in the PCA analysis since derivation is based on variance maximization while higher -order statistics are considered for non-gaussianity identification in the ICA analysis. Although explained variance is a criterion for ordering the principal components, no single criterion exists for ordering the independent components. This is the reason why in literature alternative methods that at least produce a sparse mixing matrix are considered (see Zhang *et al.* [2009]). The exact amplitude and sign of the independent components cannot be determined.

Since in the ICA analysis we try to maximize the non-gaussianity of the components, the distribution of the components are fitted better using distributions with heavier tails. In figure 1.2.3 we fit both the Normal and the Variance Gamma (see Madan and Seneta [1990b]) distribution to the first three ICs and PCs of the ten GICS. In Chapter 4 we will discuss the ICA analysis in a parametric context in order to identify the differences among the components.

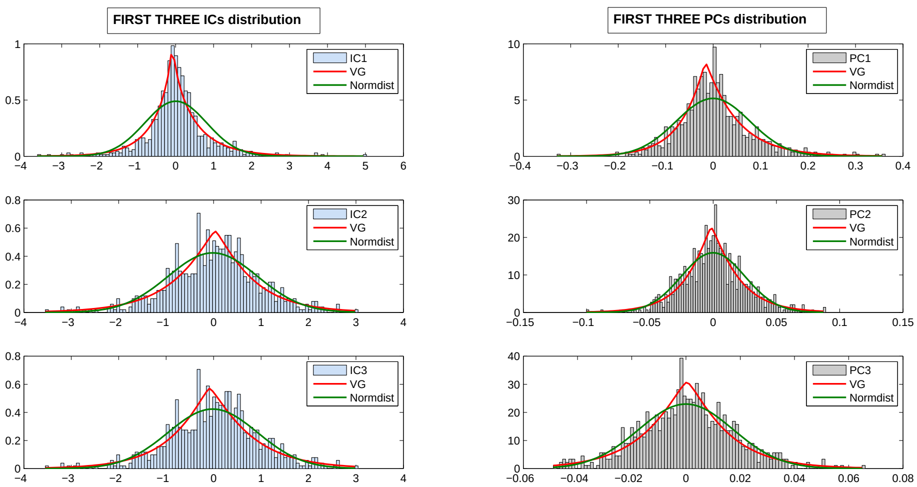


Figure 1.9: The pictures show the first three ICs and PCs of the ten GICS. The Normal and Variance distribution were fitted to each of them to evidence the differences. In particular, the first IC has a peaked distribution and the normal assumption doesn't hold.

Chapter 2

Risk Measures: VaR, CVaR and Expectiles

Risk analysis is most commonly considered in the context of a particular risk measure leading to concepts such as mean-variance optimization or diversification. The first risk measure considered was the standard deviation and one of the most well-known work based on this quantity is Markowitz [1952] which considers a risk-return optimization problem. In the last twenty years, the attention given to the correct definition of risk increased. The most used risk measure is the ValueAtRisk (VaR henceforth) RiskMetrics [1996].

Despite its popularity, VaR has two main disadvantages. The first is that it is not a coherent measure in the sense of Artzner *et al.* [1999] meaning that the diversification principle is not ensured. The second drawback is due to the fact that it does not fully take account for the tail risk. Indeed, being a point of the distribution, it does not consider the amount of loss below its value. The international regulators are thinking about moving to another risk measure for the capital requirements and, although the Expected Shortfall (ES henceforth) is a coherent risk measure, it does not seem to be the best alternative. Indeed its backtesting is more difficult to perform and, moreover, methodologies used for the estimation of the ES are not directly comparable.

In fact risk measures for which direct comparison is possible satisfy the elicitability property (see Gneiting [2011]). A statistical functional is elicitabile if it is a solution of a minimization of a suitable score function. In this context a particular attention was recently given to another statistical quantity: the expectile. Introduced in Newey and Powell [1987], it is the result of an Asymmetric Least Square regression. The expectile based risk measure seems to have nice properties.

In particular, in this chapter we review the definition of a risk measure and its mathematical properties. Then we focus on the VaR and the ES before introducing the elicitability property. An empirical investigation is performed on equity and credit risk market data for the extraction of implicit values for the weight parameter in the expectile definition.

2.1 Risk measures: mathematical properties and conformity to Basel III

Let χ be the family of the random variables defined on the same probability space $(\Omega, \mathcal{F}, P)$ describing the profit and loss (or return) of a given portfolio (or asset). The risk measure associated to $X \in \chi$ is defined as follows:

Definition 3. *A **risk measure** is a map $\rho : \chi \to \Re$ meaning that $\rho(X) \in \Re$.*

From an economic point of view $\rho(X)$ is referred to capital requirements. Depending on the sign it can have two interpretations :

- **Positive** : $\rho(X)$ is the additional amount of money to give as a margin so that the new position becomes acceptable.
- **Negative** : $\rho(X)$ is the maximum amount of money that can be withdrawn in order to maintain the acceptability of the position.

Portfolio managers use risk measures to quantify the riskiness of their position for a fixed horizon of time **T** and usually a confidence level α.
From a theoretical point of view, a very appealing class of risk measures are the coherent ones. Following the axiomatic approach introduced in Artzner *et al.* [1999], $\rho(X)$ is a coherent risk measure if it satisfies the following properties:

- **Translation Invariance** For all $\lambda \in \Re$ and for all $X \in \chi$ we have $\rho(X - \lambda) = \rho(X) - \lambda$.
- **Monotonicity** For all $X, Y \in \chi$ such that $X \leq Y$ we have $\rho(X) \leq \rho(Y)$.
- **Positive Homogeneity** For all $\lambda \geq 0$ and for all $X \in \chi$ we have $\rho(\lambda X) = \lambda\rho(X)$.
- **Subadditivity** For all $X, Y \in \chi$ we have $\rho(X + Y) \leq \rho(X) + \rho(Y)$

2.1.1 Value At Risk and Expected Shortfall

Let X be a r.v defined on the probability space $(\Omega, \mathcal{F}, P)$. $q \in \Re$ is an $\alpha-$quantile with $\alpha \in (0, 1)$ if:

$$P(X < q) \leq \alpha \leq P(X \leq q) \tag{2.1}$$

The biggest $\alpha-$quantile of X is :

$$q_\alpha^+(X) = \inf\{x \in R; F_X(x) > \alpha\} \tag{2.2}$$

while the lowest $\alpha-$quantile of X is defined as:

$$q_\alpha^-(X) = \inf\{x \in R; F_X(x) \geq \alpha\} \tag{2.3}$$

For continuous rv's we have that $q_\alpha^+(X) = q_\alpha^-(X) = q_\alpha(X)$ though we can use both for defining VaR. In particular, we define the ValueAtRisk with α level of confidence as:

$$VaR_\alpha(X) = -q_\alpha(X) \tag{2.4}$$

The notion of Value at Risk was first introduced in a technical report of J.P.Morgan in 1996. Its great success is due to the fact that it entered in the requirements of the Basel committee although it does not satisfy the subadditivity property.
Another crucial risk measure is the Expected Shortfall Acerbi [2002] defined as:

$$ES_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha VaR_\beta(X)\,d\beta \tag{2.5}$$

In the particular case where the distribution of X is continuous, Expected Shortfall coincides with the Conditional Value at Risk (CVaR) introduced in Rockafellar and Uryasev [2000]. For $\alpha \in (0,1)$ the CVaR at confidence level α is defined by:

$$CVaR_\alpha(X) = \int_{-\infty}^{+\infty} z dF_X^\alpha(z) \tag{2.6}$$

where

$$F_X^\alpha(z) = \begin{cases} 0 & z < VaR_{\alpha,t}(X) \\ \frac{F_X(z)-\alpha}{1-\alpha} & z \geq VaR_{\alpha,t}(X) \end{cases} \tag{2.7}$$

In this thesis we will use the Expected Shortfall for continuous and empirical distributions. The interpretation will be that of a conditional mean based on Tail Conditional(TCE) definition:

$$TCE_\alpha(X) = E[-X|X \leq -VaR_\alpha(X)] \tag{2.8}$$

That is, for us CVaR, ES and TCE will refer to the same quantity. As we have seen , the two considered risk measures have both Pros and Cons. In particular VaR is not subadditive, its estimators are stable and backtesting is straightforward. On the other side we have that ES is coherent, is sensitive to extreme values but its estimators are less stable than the VaR estimators as shown in Yamai and Yoshiba [2005]. Backtesting for VaR it is simple and a lot of work has been done (see Christoffersen [2010] for a general overview) while for ES is not the same since it is not an elicitable functional as explained in the next section.

2.2 Elicitability and Expectiles

2.2.1 Elicitability: forecast and comparison of estimation procedures

Savage [1971] introduced the notion of elicitability for statistical functionals. In this section we follow the presentation given in Gneiting [2011].

Let F be the distribution function of a r.v Y. Suppose we make a probabilistic forecast F for the events concerning Y. The future observations Y are the realizations of a random variable and we need a scoring method to compare the predictive performance of forecasters. There are two types of forecasts: probabilistic and point forecasts. The first considers the problem of finding any probability measure $F \in \mathcal{F}$ related to Y while in the second case we seek for a forecast of the distribution function F. For example the forecast of the mean, the mode and the quantile associated to a distribution enter in the last type of problems. We will focus on point forecast results but before explaining how all this is related to the considered risk measures let us recall some important results known in literature. Consider a family $\mathcal{F}$ of probability measures on $(I, \mathcal{B}(I))$.

Definition 4. *Statistical functional or functional*
A functional is a map $T : \mathcal{F} \to P(I)$ where $P(I)$ is the power set of the interval I.

Definition 5. *Consistent Scoring function*
A scoring function S is said to be consistent for a specified functional T defined on the the class of probability measures $\mathcal{F}$ if are satisfied both conditions:

- *$E_F S(x, Y)$ exists and it is finite $\forall x \in I$ and $\forall F \in \mathcal{F}$.*
- *$E_F S(t, Y) \leq E_F S(x, Y)$ for all $x \in I$, $\forall F \in \mathcal{F}$ and $\forall t \in T(F)$.*

where the scoring function is usually an error measure.

Definition 6. *Strictly Consistent Scoring function*
A scoring function S is said to be strictly consistent for a specified functional T defined on the the class of probability measures $\mathcal{F}$ if:

- *$E_F S(t, Y) = E_F S(x, Y)$ $\forall x \in I$, $\forall F \in \mathcal{F}$ and $\forall t \in T(F)$.*

The equality of the expectations ensures that $t \in T(F)$ implies $x \in T(F)$.

Definition 7. *Elicitable functional*
A functional T is elicitable if exists a scoring function S that is strictly consistent for it.

Recently, Bellini and Bignozzi [2013] suggest a slightly more restrictive definition more suited for financial applications by adding the convexity requirement for the scoring function.
The following result is useful in the case we want to show that a given functional is not elicitable: Osband [1985] shows that the level sets of an elicitable functional are convex. This property requires that if $F_0 \in \mathcal{F}$, $F_1 \in \mathcal{F}$ and for $p \in (0, 1)$ $F_p = (1-p)F_0 + pF_1 \in \mathcal{F}$ the conditions $t \in T(F_0)$ and $t \in T(F_1)$ imply that $t \in (F_p)$. In fact, if T is elicitable, meaning that it exists a strictly consistent scoring function S, the linearity property of expectations and the elicitability definition, for any $x \in I$ suggest that:

$$E_{F_p} S(t, Y) = pE_{F_0} S(t, Y) + (1-p)E_{F_1} S(t, Y) \leq pE_{F_0} S(x, Y) + (1-p)E_{F_1} S(x, Y) = E_{F_p} S(x, Y) \tag{2.9}$$

ensuring that $t \in (F_p)$.
For a given functional T, in order to get precise proofs, we need some smoothness conditions on the scoring function S. A common choice is the class of Generalized Piecewise linear functions of order α:

Definition 8. *Generalized Piecewise linear (GPL) of order* α
A function $S_g : I \times I \to [0, +\infty)$ *of the form:*

$$S_g(x, y) = (1-\alpha)[g(x) - g(y)]1_{y \leq x} + \alpha[g(x) - g(y)]1_{y\, x} \tag{2.10}$$

with $g : I \to \Re$ *being a non decreasing function.*

In the particular case where g is a linear function, S_g coincides with the asymmetric piecewise loss function used in quantile regression. The function S_g, for g linear, is a strictly consistent scoring function for the functional $q_\alpha(Y) = F_Y^{-1}(\alpha)$. It is simple now to understand why VaR is an elicitable functional. This property ensures that exists competing alternatives that are directly comparable which to measure the error we make by assuming a particular distribution for the data considered. The sufficiency condition for elicitability doesn't depend on the smoothness conditions for the GPL function. The CVaR functional is not elicitable since the level set of the functionals are not convex. In Gneiting [2011] a counterexample for the class $\mathcal{F}$ containing measures with finite support . In particular suppose $a, b, c, d \in I$ with $a < b < c < 0.5(b + d)$ and define two probability measures:

$$F_1 = \alpha\delta_a + 0.5(1-\alpha)[\delta_b + \delta_d]F_2 = \alpha\delta_c + (1-\alpha)\delta_{(b+d)/2} \tag{2.11}$$

It is easy to observe that for $\alpha \leq 1/3$:

$$CVaR_\alpha(F_1) = CVaR_\alpha(F_1) = 0.5(b + d) \tag{2.12}$$

CVaR doesn't have convex level sets since:

$$CVaR_\alpha(0.5(F_1 + F_2)) = 0.25(b + c + 2d) > 0.5(b + d) \tag{2.13}$$

The CVaR associated to the convex combination of the probability measures is higher than the weighted sum of the single CVaRs since $b < d$.
The BIS Consultative Document BIS [2012] poses the question of what are the likely operational constraints with moving from VaR to ES, including any challenges in delivering robust backtesting, and how might these be best overcome. The absence of elicitability property probably will justify the embargo for the use of the ES for capital requirements.

2.2.2 Expectiles: coeherent and elicitable

For $\tau \in (0,1)$, the τ-expectile of a r.v X with finite mean is defined as:

$$e_\tau(X) = argmin_{m\in\Re} E[\tau max(m-X,0)^2 + (1-\tau)max(X-m,0)^2] \tag{2.14}$$

It can easily be shown that $e_\tau(X)$ satisfies the relation:

$$e_\tau(X) = \frac{\tau E\left[XI_{\{X\leq e_\tau(X)\}}\right] + (1-\tau)E\left[XI_{\{X>e_\tau(X)\}}\right]}{\tau P[X\leq e_\tau(X)] + (1-\tau)P[X>e_\tau(X)]} \tag{2.15}$$

The risk measure associated to the expectile is:

$$\rho(X) = -e_\tau(X) \tag{2.16}$$

The main characteristics that make it a candidate as a risk measure are the fact that it is homogenous and coherent for particular choice of τ (see Bellini *et al.* [2013]). Under the condition of finiteness for the second moment of F, the τ-expectile is the optimal point forecast if the scoring function is:

$$S_\tau(x,y) = |1_{x\geq y} - \tau|\,(x-y)^2 \tag{2.17}$$

Bellini and Bignozzi [2013] showed that expectiles are the only coherent risk measures to be elicitable it exists a clear methodology to perform backtesting. The alternative procedures can be ranked by using the scoring function since if the scoring function $S(x,y)$ is consistent for the considered risk measure $\rho(X)$, then the averaged realized score:

$$\frac{1}{n} = \sum_{i=1}^{n} S(\rho_i, x_i) \tag{2.18}$$

should be small.
The main difficulty in the use of the expectile based risk measure is the choice of the parameter τ. The arbitrary choice for the parameter value can be faced by recalling that the expectile parameter τ and quantiles are connected since (see Jones [1994]) :

$$\tau = \frac{\alpha q_\alpha - \int_{-\infty}^{q_\alpha} x dF(x)}{E[X] - 2\int_{-\infty}^{q_\alpha} x dF(x) - (1-2\alpha)q_\alpha} \tag{2.19}$$

It is though possible, once fixed α, to obtain a value for τ for the given distribution of the data considered.

In order to have a better intuition of the parameters appearing in the quantile and expectile definitions, we perform a naive backtesting analysis. The $S\&P500$ returns for the time-period

12/09/2009- 12/09/2012 are modeled as as:

$$r_t = \sigma_t \epsilon_t \tag{2.20}$$

and following Bollerslev [1986] we suppose $\sigma_t \sim GARCH(1,1)$, that is:

$$\sigma_t^2 = \omega + \beta_1 r_{t-1}^2 + \beta_2 \sigma_{t-1}^2 \tag{2.21}$$

where β_1 and β_2 are the model parameters. From the GARCH(1,1) fitting, we get the sequence $\hat{\sigma}_t$ $t = 1, ..., n$. Consider :

$$\hat{\epsilon}_t = \frac{r_t}{\hat{\sigma}_t} \tag{2.22}$$

and compute the expectile of this sequence for fixed τ as the solution of an Asymmetric Least Square Regression :

$$e_\tau = argmin_m \frac{1}{n} \sum_{i=1}^{n} \left[\tau(\hat{\epsilon}_t - m)^2 I_{\hat{\epsilon}_t \geq m} + (1-\tau)(m - \hat{\epsilon}_t)^2 I_{\hat{\epsilon}_t < m} \right] \tag{2.23}$$

As a first attempt, we fix two values for τ respectively 0.05 and 0.01. The expectiles obtained daily for these two parameter values are plotted together with the α- quantile for $\alpha = 0.05$. The

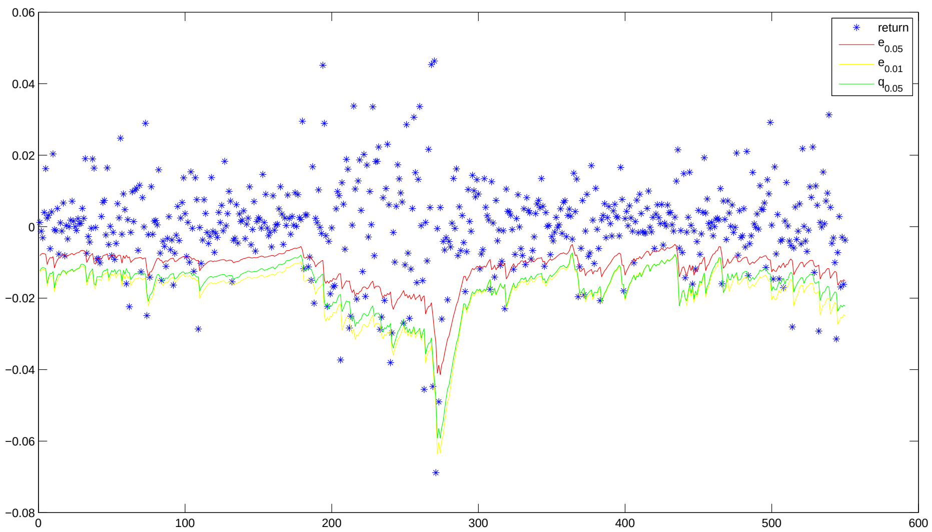


Figure 2.1: The $S\&P500$ returns for the time-period 12/09/2009- 12/09/2012 is modeled using a GARCH(1,1). The daily forecast for the quantities $e_{0.05}$, $e_{0.01}$ and $q_{0.05}$ is then compared with the actual returns in order to compute the exceedance rate.

time frame considered for the estimation contains 200 observations. Daily we check whether the actual return observed is less than the expectile or quantile estimated in the previous step. The returns for which this happen form the sequence of exceedances or hittings. The exceedance rate is defined as the fraction of the exceedances over the total number of observations. We compute $e_{0.01}$ and obtain an exceedance rate of 4.36% while for $e_{0.05}$ the exceedance rate is 12.55%. The exceedance rate for $VaR_{0.05}$ is 5.45%. From this simple exercise we confirm the fact that while α can be interpreted as the probability of observing returns lower than the corresponding VaR_α value, for τ this is not possible. In fact, it has been interpreted as the prudential parameter in Kuan *et al.* [2009b] since it the weight given to the tails in the ALS regression problem. Now we perform an empirical analysis based on the expectile-quantile relation given in equation 2.19. Fixed α, we use the empirical distribution to compute the value for τ. First we apply our analysis to the S&P500 and to the ten GICS return distributions. Then we move on to the credit market and consider the daily z-spread variations for eleven European countries.

From the empirical distribution of the daily log returns for the S&P500 index, we get the τ for α values that are usually used in finance: $0.01, 0.05$ and 0.10.

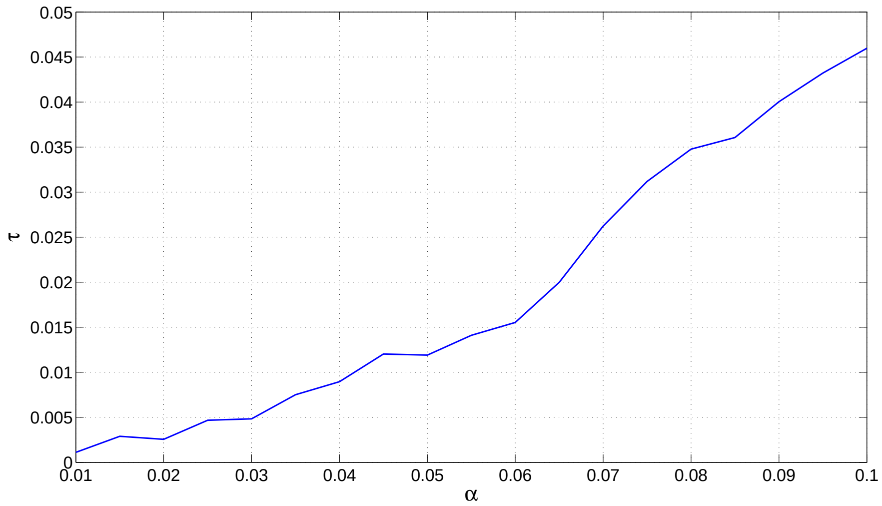


Figure 2.2: In this figure we plot the τ for different α values using the empirical distributions of the S&P500 index. The values were obtained from the formula that relates α to τ.

α	0.01	0.05	0.1
τ	0.0011	0.0119	0.0459
q_α	-0.0226	-0.0159	-0.0107
e_τ	-0.0227	-0.0161	-0.0108

Table 2.1: In this table we give the τ values obtained using the corresponding empirical distribution for the S&P500 index when α is $0.01, 0.05$ and 0.10. The corresponding empirical expectile and quantile values are given and we can see they differ quite slightly.

For the same α values, we repeat the analysis for the ten GICS sector indexes.

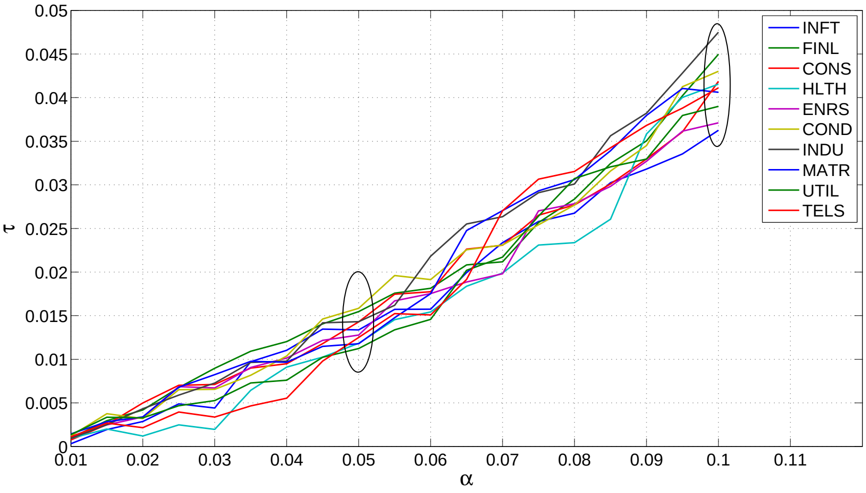


Figure 2.3: In this figure we plot τ values for different α using the empirical distributions of the ten GICS. Observe that for $\alpha = 0.05$ the ten sectors τ-s are more concentrated than the τ-s obtained for $\alpha = 0.10$.

α		0.01	0.05	0.1
τ	INFT	0.0003	0.0134	0.0363
	FINL	0.0010	0.0155	0.0450
	CONS	0.0008	0.0143	0.0419
	HLTH	0.0009	0.0118	0.0415
	ENRS	0.0011	0.0128	0.0371
	COND	0.0012	0.0158	0.0430
	INDU	0.0009	0.0143	0.0475
	MATR	0.0014	0.0118	0.0406
	UTIL	0.0014	0.0112	0.0390
	TELS	0.0011	0.0125	0.0411

Table 2.2: In this table we give the τ value obtained using the corresponding empirical distribution for the ten GICS when α is $0.01, 0.05$ or 0.10.

Observe that for higher α values, the set of implicit values for τ becomes larger. However, since the distributions of the ten GICS sector indexes are similar, the corresponding τ-s are similar too.

Let us consider the dataset composed by the daily variations of the z-spreads for eleven sovereigns in the euro zone going from 20/10/2010 to 20/10/2012. With respect to the equity dataset, now there in bigger variability in the τ values for each fixed α. Moreover, we systematically observe higher values for τ than those obtained for the same α values when the GICS dataset is considered.

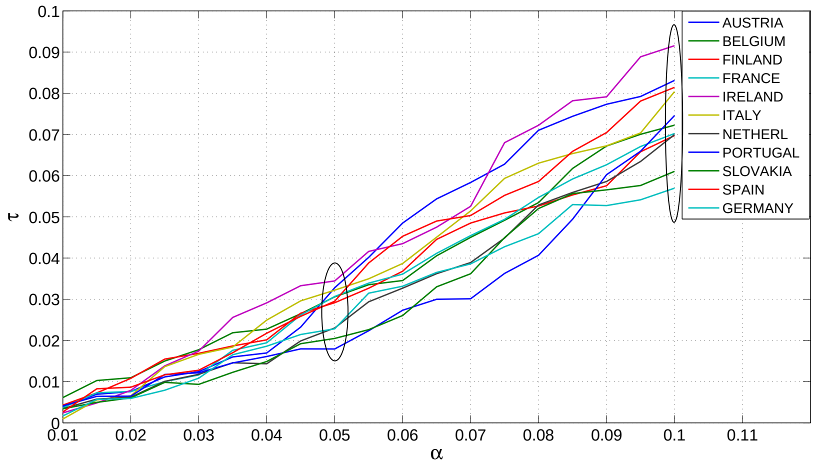


Figure 2.4: In this figure we show τ for different α using the empirical distributions of eleven European sovereign z-spreads.

α		0.01	0.05	0.1
τ	Austria	0.0039	0.0328	0.0831
	Belgium	0.0061	0.0306	0.0723
	Finland	0.0043	0.0292	0.0697
	France	0.0036	0.0307	0.0702
	Ireland	0.0024	0.0344	0.0915
	Italy	0.0009	0.0322	0.0804
	Netherl	0.0034	0.0230	0.0700
	Portugal	0.0041	0.0179	0.0746
	Slovakia	0.0034	0.0205	0.0610
	Spain	0.0026	0.0295	0.0814
	Germany	0.0017	0.0228	0.0570

Table 2.3: In this table we give the τ value obtained using the corresponding empirical distribution for the sovereign z-spread daily variations when α is $0.01, 0.05$ or 0.10.

The difference in the results suggests that for fixed α values, the implicit τ depend on the shape of the dataset distribution. In fact, if we compute the kurtosis level for the z-spread dataset, the value is much higher than when we consider equity indexes. The results show that if we want to use an expectile based risk measure, the parameter τ should not be chosen arbitrarily but it it should be dependent on some statistics of the distribution considered. From the two exercises performed, it seems that τ can be defined as an increasing function of the kurtosis level. Further investigations should be performed in order to model the relation between them.

	INFT	FINL	CONS	HLTH	ENRS	COND	INDU	MATR	UTIL	TELS
kurtosis	3.6637	4.1544	3.4976	3.8195	3.5810	3.9228	3.6497	3.6099	3.2496	3.1990

Table 2.4: Kurtosis level for the empirical distribution of the daily log returns for the ten GICS indexes.

	Austria	Belgium	Finland	France	Ireland	Italy	Netherl	Portugal	Slovakia	Spain	Germany
kurtosis	10.3713	12.9797	10.8325	9.3513	12.6521	8.2672	9.4645	20.8646	5.6186	7.6921	6.3067

Table 2.5: Kurtosis level for the empirical distribution of the sovereign z-spread daily variation for eleven countries. The departure from the normal distribution is evident

2.3 Risk decomposition

We cannot define universally the most appropriate financial risk measure. Users have to make choices of the properties of risk measures that are most important for their purposes. Let us see how the three considered risk measures can be used in risk decomposition. The homogeneity property results fundamental.

2.3.1 Homogeneity and Euler decomposition

A function $f(x, y)$ is said to be homogeneous of order n if:

$$f(\lambda x, \lambda y) = \lambda^n f(x, y) \tag{2.24}$$

For homogeneous functions holds the Euler theorem:

$$\frac{\partial f(x, y)}{\partial x} x + \frac{\partial f(x, y)}{\partial y} y = n f(x, y) \tag{2.25}$$

This theorem can be generalized for more variables and it is the basis for the risk contribution methodology introduced in Tasche [1999]
Consider a linear model relating portfolio returns π with factors in F:

$$\pi = \beta F \tag{2.26}$$

Any homogeneous risk measure $R(\pi)$ can be written as the sum:

$$R(\pi) = \sum_{i=1}^{n} \beta_i \frac{\partial R(\pi)}{\partial \beta_i} = \sum_{i=1}^{n} RC_i \tag{2.27}$$

where the risk contribution of the i-th risk factor is:

$$RC_i(\pi) = \beta_i \frac{\partial R(\pi)}{\partial \beta_i} \tag{2.28}$$

This result has been widely used in finance since it gives a simple rule for attributing risk (see for example Meucci [2007] and Marchioro and Borrello [2013]).

2.3.2 Risk measure decomposition formulas

Sharpe [2002] affirms that a mere mathematical decomposition of portfolio risk can not necessarily be seen as risk contribution. However, this mathematical decomposition of portfolio risk has a financial meaning. Consider the following risk measures :

Value at Risk

$$VaR_\alpha(\pi) = -q_\alpha^+(\pi) = -inf\,\{x \in \Re | F_\pi(x) > \alpha\} \tag{2.29}$$

Expected Shortfall (or CVaR)

$$ES_\alpha(\pi) = -E\left[\pi\,|\pi \leq -VaR_\alpha(\pi)\right] \tag{2.30}$$

Expectile based risk measure

$$\rho_\tau(\pi) = -\frac{\tau E\left[\pi I_{\{\pi \leq e_\tau(\pi)\}}\right] + (1-\tau)E\left[\pi I_{\{\pi > e_\tau(\pi)\}}\right]}{\tau P[\pi \leq e_\tau(\pi)] + (1-\tau)P[\pi > e_\tau(\pi)]} \tag{2.31}$$

and the corresponding risk contributions based on the:

Value-at-risk (see Gourieoux *et al.* [2000])

$$RC_i = -E\left[F_i\,|\pi = VaR_\alpha(\pi)\right]\beta_i \tag{2.32}$$

Expected Shortfall (see Scaillet [2002])

$$RC_i = -E\left[F_i\,|\pi \leq -VaR_\alpha(\pi)\right]\beta_i \tag{2.33}$$

Expectiles (see Tasche (2013)**)**

$$RC_i = -\frac{\tau E\left[F_i I_{\{\pi \leq e_\tau(\pi)\}}\right] + (1-\tau)E\left[F_i I_{\{\pi > e_\tau(\pi)\}}\right]}{\tau P[\pi \leq e_\tau(\pi)] + (1-\tau)P[\pi > e_\tau(\pi)]}\beta_i \tag{2.34}$$

The formulas described are easy to implement. In fact, we will perform risk attribution considering the three described risk measures in the following sections.

2.3.3 Empirical results

Here, we consider the data composed by daily log returns from 11/02/2010 to 11/02/2013 on the SP500 and the ten GICS. The ES at level $\alpha = 0.05$ and the expectile for $\tau = 0.012$ were computed for the general index and for each sector. The portion of risk associated to each sector seems to remain stable going from one measure to the other.

	SPX	**INFT**	**FINL**	**CONS**	**HLTH**	**ENRS**	**COND**	**INDU**	**MATR**	**UTIL**	**TELS**	**Residual**
ES_0.05	0.0191	0.0037	0.0037	0.0015	0.0015	0.0024	0.0022	0.0024	0.0008	0.0004	0.0003	2.72E-05
		19.6%	19.3%	8.0%	8.1%	12.7%	11.7%	12.8%	4.3%	1.9%	1.5%	0.1%
e_0.012	0.0153	0.0032	0.0028	0.0013	0.0010	0.0021	0.0017	0.0019	0.0007	0.0003	0.0002	2.70E-05
		20.8%	18.6%	8.8%	6.6%	13.9%	11.2%	12.7%	4.4%	1.7%	1.0%	0.2%

Table 2.6: Risk decomposition when the factors are the GICS

Chapter 3

User defined risk factors

In the previous chapters we discussed the problem of identifying the factors that can be useful in explaining the overall return/risk. As we already saw, the considered factors can have a financial interpretation or can be the result of a mathematical procedure though interpretation is not straightforward. In this chapter we face the problem of attributing risk to factors that are directly chosen by the investor and different from those that fully determine the total portfolio/asset returns.

Now we suppose that the explicative factors have already been identified. We want to quantify the influence of a financial variable that doesn't enter in the set of risk factors considered. For example, consider a portfolio of fixed rate bonds whose risk factors are the changes in value of zero rates and credit spreads and suppose that an abrupt movement in the Oil index price is observed. Though not considered before, the exposure of the portfolio to the Oil index becomes relevant. Since this variable is not uncorrelated with the initial risk factors, its inclusion in the risk decomposition model requires some efforts in order to avoid the creation of possible collinear factors. This problem has become particularly relevant especially since the recent financial crisis has suggested that market variables are more related to each other than one could think. Meucci [2007] considers the problem in the context where the new defined factors are uncorrelated from the residuals and linear transformations of the original risk factors. We extend the analysis by attributing risk to the ICs since the conditions posed in the approach are satisfied in the ICA analysis.

However, in order to be more realistic, we study the problem in the case of non linear transformations. Menchero and Poduri [2008] propose a model that is justified by the need of attributing risk to the allocation decisions made by the portfolio manager. We consider the question from a different point of view. After the allocation decision has been made, we define some new factors uncorrelated with the user defined factors. They are the projections of the original factors on the space orthogonal to that generated from the user defined factors. In this way we are able to decompose the return vector in three uncorrelated quantities: the user defined factors, the projected factors and the idiosyncratic error term. The model proposed is easy to implement and each risk component has a financial interpretation.

3.1 Risk Contribution from user defined factors

In this section we discuss the problem of attributing risk to user defined risk factors following the approach introduced in Meucci [2007]. We consider the return vector $r_{1\times t} = \beta F$ that can be attributed to n factors:

$$F_{n\times t} = \begin{bmatrix} F_1 \\ F_2 \\ ... \\ F_i \\ ... \\ F_n \end{bmatrix} = \begin{bmatrix} F_{1,1} & F_{1,2} & ... & F_{1,t} \\ F_{2,1} & F_{2,2} & ... & F_{2,t} \\ ... & ... & ... & ... \\ F_{i,1} & F_{i,2} & ... & F_{i,t} \\ ... & ... & ... & ... \\ F_{n,1} & F_{n,2} & ... & F_{n,t} \end{bmatrix} \tag{3.1}$$

with exposures given by the vector:

$$\beta_{1\times n} = \begin{bmatrix} \beta_1 & \beta_2 & ... & \beta_i & ... & \beta_n \end{bmatrix} \tag{3.2}$$

The risk associated to the asset is the weighted sum of each factor marginal risk:

$$R(\beta) = \sum_{i=1}^{n} \beta_i \frac{\partial R(\beta)}{\partial \beta_i}. \tag{3.3}$$

As we have seen in Chapter 2, if we choose to consider the expected shortfall as the risk measure:

$$R(\beta) = E\left[-\beta F \,|r \leq \; -VaR_\alpha(\beta F)\right] \tag{3.4}$$

the marginal contribution to risk associated to the i-th factor is:

$$\frac{\partial R(\beta)}{\partial \beta_i} = -E\left[F_i \,|r \leq \; -VaR_\alpha(\beta F)\right] \tag{3.5}$$

Consider now k new factors $\tilde{F}$, with $k \leq n$, obtained as a linear transformation of the original factors through matrix $P_{k\times n}$:

$$\tilde{F}_{k\times t} = P_{k\times n} F_{n\times t} = \begin{bmatrix} P_{1,1} & P_{1,2} & ... & P_{1,n} \\ P_{2,1} & P_{2,2} & ... & P_{2,n} \\ ... & ... & ... & ... \\ P_{k,1} & P_{k,2} & ... & P_{k,n} \end{bmatrix} \begin{bmatrix} F_{1,1} & F_{1,2} & ... & F_{1,t} \\ F_{2,1} & F_{2,2} & ... & F_{2,t} \\ ... & ... & ... & ... \\ F_{i,1} & F_{i,2} & ... & F_{i,t} \\ ... & ... & ... & ... \\ F_{n,1} & F_{n,2} & ... & F_{n,t} \end{bmatrix} \tag{3.6}$$

The new defined factors $\tilde{F}$ don't span the entire market and an error term appears:

$$r = \tilde{\beta}\tilde{F} + \epsilon \tag{3.7}$$

The new exposures $\tilde{\beta}$ need to be defined following any criteria. In a linear regression problem the exposures are chosen such that the new variables (here $\tilde{F}$) are uncorrelated with the residuals:

$$Cov(\tilde{F}, \epsilon') = 0 \tag{3.8}$$

Substitute the return definition in terms of the original factors F :

$$Cov(\tilde{F}, r' - \tilde{F}'\tilde{\beta}') = 0 \tag{3.9}$$

and after performing some linear algebra passages we have:

$$Cov(PF, F'\beta') - Cov(PF, F'P')\tilde{\beta}' = 0 \tag{3.10}$$

The exposures to the new factors are:

$$\tilde{\beta}' = [P\Sigma_F P']^{-1} P\Sigma_F \beta' \tag{3.11}$$

where Σ_F is the covariance matrix of the original set of risk factors. Given the rule that identifies the exposures $\tilde{\beta}$, we can compute the risk contribution for the new factors. The part of r that is not attributed to the new factors $\tilde{F}_{k\times t} = P_{k\times n}F_{n\times t}$ is attributed to the remaining theoretical factors $\hat{F}$, i.e :

$$r = \tilde{\beta}\tilde{F} + \hat{\beta}\hat{F} \tag{3.12}$$

$$\hat{F}_{(n-k)\times t} = \hat{P}_{(n-k)\times n}F_{n\times t} \tag{3.13}$$

where $\hat{P}_{(n-k)\times n}$ is such that the matrix:

$$\bar{P}_{n\times n} = \begin{bmatrix} P \\ \hat{P} \end{bmatrix} \tag{3.14}$$

is invertible. The portfolio return vector can be written in the form :

$$r = \beta F = \beta \bar{P}^{-1}\bar{P}F = \begin{bmatrix} \tilde{\beta} & \hat{\beta} \end{bmatrix} \begin{bmatrix} \tilde{F} \\ \hat{F} \end{bmatrix} \tag{3.15}$$

The total risk using the Euler rule is:

$$R = \tilde{\beta}\frac{\partial R}{\partial \tilde{\beta}} + \hat{\beta}\frac{\partial R}{\partial \hat{\beta}} = \begin{bmatrix} \tilde{\beta}_1 & ... & \tilde{\beta}_k \end{bmatrix} \begin{bmatrix} \frac{\partial R}{\partial \tilde{\beta}_1} \\ ... \\ ... \\ \frac{\partial R}{\partial \tilde{\beta}_k} \end{bmatrix} + \begin{bmatrix} \hat{\beta}_1 & ... & \hat{\beta}_{(n-k)} \end{bmatrix} \begin{bmatrix} \frac{\partial R}{\partial \hat{\beta}_1} \\ ... \\ ... \\ \frac{\partial R}{\partial \hat{\beta}_{(n-k)}} \end{bmatrix} \tag{3.16}$$

The complete vector of exposures is:

$$\begin{bmatrix} \tilde{\beta} & \hat{\beta} \end{bmatrix} = \beta \bar{P}^{-1} \tag{3.17}$$

from where we have:

$$\begin{bmatrix} \tilde{\beta} & \hat{\beta} \end{bmatrix} \begin{bmatrix} P \\ \hat{P} \end{bmatrix} = \beta \tag{3.18}$$

In order to get the marginal exposures we need to express β as a linear transformation of $\tilde{\beta}$ making clear the relation between them:

$$\beta = \tilde{\beta}P + \hat{\beta}\hat{P} \tag{3.19}$$

Recall the differentiation rule

$$\frac{\partial XA}{\partial X} = A \tag{3.20}$$

and the chain rule applied to matrices and get that :

$$\frac{\partial R}{\partial \tilde{\beta}} = \begin{bmatrix} \frac{\partial R}{\partial \tilde{\beta}_{1,1}} & \frac{\partial R}{\partial \tilde{\beta}_{1,2}} & ... & \frac{\partial R}{\partial \tilde{\beta}_{1,k}} \end{bmatrix} = \begin{bmatrix} \frac{\partial R}{\partial \beta_{1,1}} & \frac{\partial R}{\partial \beta_{1,2}} & ... & \frac{\partial R}{\partial \beta_{1,k}} \end{bmatrix} \begin{bmatrix} \frac{\partial \beta_{1,1}}{\partial \tilde{\beta}_{1,1}} & \frac{\partial \beta_{1,1}}{\partial \tilde{\beta}_{1,2}} & ... & \frac{\partial \beta_{1,1}}{\partial \tilde{\beta}_{1,k}} \\ ... & ... & ... & ... \\ \frac{\partial \beta_{1,n}}{\partial \tilde{\beta}_{1,1}} & \frac{\partial \beta_{1,n}}{\partial \tilde{\beta}_{1,2}} & ... & \frac{\partial \beta_{1,n}}{\partial \tilde{\beta}_{1,k}} \end{bmatrix} \tag{3.21}$$

The contribution to risk of each new factor becomes the product of the initial contribution to risk with the transposed matrix used for the factor transformation:

$$\frac{\partial R}{\partial \tilde{\beta}} = \frac{\partial R}{\partial \beta} P' \tag{3.22}$$

This approach requires only the knowledge of the transformation matrix but it can be applied in the cases when the new considered factors are uncorrelated and linear transformations of the original risk factors. In his paper Meucci [2007] performs an empirical analysis using the PCs as new uncorrelated factors. We go on and use this approach for decomposing risk when the factors considered are the ICs. The independence of the ICs ensures their uncorrelation, though the conditions required are satisfied.

3.1.1 Risk contribution using PCA and ICA

Let us consider the same dataset as in Section 2.5, i.e we consider the $S\&P500$ index as a portfolio whose returns depend on the initial risk factors, the GICS sector indexes. First, we investigate the risk attributed to the 7 PCs that explain at least 98% of the portfolio total variability. The risk measures considered are the Expected Shortfall for $\alpha = 0.05$ and the expectile for $\tau = 0.012$. As we expected, the first PC contribution to the portfolio risk is substantial for both the risk measures considered.

		PC1	**PC2**	**PC3**	**PC4**	**PC5**	**PC6**	**PC7**	**Residual**
ES_0.05	0.0191	0.0192	0.0004	-0.0003	-4.97E-05	-3.69E-06	-1.64E-05	-5.93E-05	-3.47E-05
		100.6%	1.9%	-1.6%	-0.3%	0.0%	-0.1%	-0.3%	-0.2%
e_0.012	0.0153	0.0151	0.0004	-0.0002	1.92E-05	-8.27E-06	-1.96E-05	0.0001	1.41E-05
		98.5%	2.4%	-1.6%	0.1%	-0.1%	-0.1%	0.7%	0.1%

Table 3.1: Risk decomposition using PCs as user defined factors. The analysis is performed considering on the $S\&P500$ and considering as initial risk factors the GICS sector indexes.

When the factors considered are the ICs, it is not possible to order them according to the variance explained. Being in front of an open problem, we decide to consider the same number of factors as suggested in the PCA analysis. That is, we consider the first 7 ICs as the sources of randomness in the market while the remaining 3 are grouped together in the residual term. Observe that, as we supposed, the contribution to risk of the ICs does not descrease systematically as we go on with the components. The risk attribution to the residual term is higher than that of the seventh IC. This suggests that further investigation should be performed on the residual term in order to extract any other IC that influences directly the portfolio risk.
Since the distribution of the statistical factors can be quite different from those of the initial risk

		IC1	**IC2**	**IC3**	**IC4**	**IC5**	**IC6**	**IC7**	**Residual**
ES_0.05	0.0191	0.0189	-0.0048	-0.0004	0.0005	0.0005	0.0036	0.0002	0.0006
		98.9%	-25.0%	-2.4%	2.6%	2.9%	18.8%	1.2%	2.9%
e_0.012	0.0153	0.0152	-0.0039	-0.0004	0.0003	0.0005	0.0031	0.0001	0.0003
		99.6%	-25.7%	-2.6%	2.3%	3.4%	20.1%	1.0%	1.9%

Table 3.2: Risk decomposition using ICs as user defined risk factors. The analysis is performed considering on the SPX and considering as initial risk factors the GICS sector indexes.

factors, the sign of the contribution to the total risk can be negative. In the next section we will attribute risk to user defined risk factors that are based on observed financial quantities and we expect that the sign of the contribution to risk for each factor to be coherent with the sign of the correlation coefficient of the portfolio and the factor considered.

3.2 Custom Factor Attribution

An intuitive attribution of risk requires that the sources of risk to be aligned with the sources of return. Improper use of return sources may lead to a number of problems like mismatch between risk and return sources, non-intuitive marginal contributions to risk or wrongly naming aggressive positions as risk reducing. Risk forecasting is an ex-ante analysis. Generally a portfolio manager knows the factor exposures β but factor return distributions F must be predicted. During the performance attribution process, that is an ex-post analysis, β and F are known at the end. The discussion in this section is about risk forecasting, i.e an ex-ante analysis and based on the methodology introduced in Menchero and Poduri [2008].
The starting point is the return attribution model. Suppose that each asset return depends on K factors and on an idiosyncratic component:

$$r = \beta F + \epsilon^{orig} \tag{3.23}$$

where the idiosyncratic returns ϵ^{orig} are mutually uncorrelated and uncorrelated with factor returns. In particular, we have that r is the $1 \times T$ vector of asset returns, β is the $1 \times K$ vector containing the weights generated from the allocation decisions, F is the $K \times T$ matrix of factor returns and ϵ is the $1 \times T$ vector of residuals. Factors F can be used for describing returns but if they don't match the manager investment process may be wrong for performance attribution. Let F^Y be the custom factors such that they reflect the investment process:

$$r = \beta^Y F^Y + \epsilon^Y \tag{3.24}$$

with the residuals ϵ^Y being neither mutually uncorrelated nor uncorrelated with factors F^Y. Most of the time factors F^Y are not able to explain all the risk but there is some interest in understanding the part of return/risk attributed to a specific decision on investing in a particular asset. In terms factors F and F^Y the portfolio return becomes :

$$r = \beta^Y F^Y + \beta F + \epsilon^{Y,orig} \tag{3.25}$$

The possible linear dependence among β^Y and β makes difficult the estimation used by regression. To avoid collinear β and β^Y create the exposures to residual factors $\tilde{\beta}$ such that:

$$\beta^{Y'} \tilde{\beta} = 0 \tag{3.26}$$

We want to decompose total risk in contributions coming from custom and residual factors. The projection operator which preserves only the components perpendicular to β^Y is :

$$P_{\beta^{Y\perp}} = I - \beta^Y ({\beta'}^Y \beta^Y)^{-1} {\beta'}^Y \tag{3.27}$$

and is used to obtain the exposures to the residual factors :

$$\tilde{\beta} = P_{\beta^{Y\perp}}\beta \tag{3.28}$$

The projection operator which preserves only the components within the space of β^Y is :

$$P_{\beta^Y} = \beta^Y({\beta'}^Y\beta^Y)^{-1}{\beta'}^Y \tag{3.29}$$

Observe that :

$$P_{\beta^Y} + P_{\beta^{Y\perp}} = I \tag{3.30}$$

and obtain :

$$r = \beta^Y F^Y + \tilde{\beta}\tilde{F} + \tilde{\epsilon} \tag{3.31}$$

where F^Y is the $L \times T$ vector of custom factors, $\tilde{F}$ is the $K \times T$ matrix of residual factors, $\tilde{\epsilon}$ is the $1 \times T$ vector of idiosyncratic returns, β^Y is the $1 \times L$ vector of custom factor exposures and $\tilde{\beta}$ is the $1 \times K$ vector of residual factor exposures. Through OLS regression factor returns estimation are :

$${F_{ols}}^Y = ({\beta'}^Y\beta^Y)^{-1}{\beta'}^Y r \tag{3.32}$$

$$\tilde{F}_{ols} = (\tilde{\beta}'\tilde{\beta})^{-1}\tilde{\beta}'r \tag{3.33}$$

If we think at r as the matrix of the returns of n assets[1] and consider the weight $n \times 1$ vector w, the portfolio return π, defined as :

$$\pi = w'r \tag{3.34}$$

becomes the sum of returns attributable to custom, residual factors and idiosyncratic factors :

$$\pi = \pi_Y + \pi_{\tilde{X}} + \pi_{\tilde{\epsilon}} \tag{3.35}$$

where $\pi_Y = w'P_Y r$, $\pi_{\tilde{F}} = w'P_{\tilde{F}}r$ and $\pi_{\tilde{\epsilon}} = w'(I_N - P_Y - P_{\tilde{F}})r$.

3.3 Custom and projected factors

3.3.1 Decomposing the residual

In this section we consider ex-post analysis, in the sense that once the factor returns are observed we can perform risk attribution to the factors. The approach is similar to that discussed in Menchero and Poduri [2008] but with some differences. As usual, we start with the hypothesis that the return can linearly be decomposed using n factors that we call F^X. The problem of attributing risk to a new set of factors in F^Y and contemporary to the old factors F^X is faced at first by considering a two step-regression. Here we describe the main steps. Given the new factors F^Y we determine

[1] r is an $n \times t$ matrix

the exposures to them by :

$$\beta^Y = (F^{Y'} F^Y)^{-1} F^{Y'} r \tag{3.36}$$

Denoting the new residuals with δ we have the decomposition of the returns in terms of the new factors and the residual term:

$$r = \beta^Y F^Y + \delta \tag{3.37}$$

Since F^Y is an arbitrary chosen factor, its contribution to the portfolio return/risk can be quite limited. In order to extract more information from the δ term, we perform another regression analysis this time using as regressors the original factors F^X. The relation of δ with factors F^X is:

$$\delta = \beta^{Projx} F^X + \hat{\epsilon} \tag{3.38}$$

where :

$$\beta^{Projx} = (F^{X'} F^X)^{-1} F^{X'} \delta \tag{3.39}$$

The new decomposition is:

$$r = \beta^Y F^Y + \beta^{Projx} F^X + \hat{\epsilon} \tag{3.40}$$

but remains the problem that F^Y and F^X could be linearly dependent.
We want to attribute risk to factors F^Y but in the same time the contribution of the idiosyncratic term must be extracted form the first regression step. That is, we aim at having a model of the form :

$$r = \beta^Y F^Y + \beta^{\tilde{X}} F^{\tilde{X}} + \epsilon^i \tag{3.41}$$

where e^i is the idiosyncratic residual uncorrelated with the custom factors F^Y and $F^{\tilde{X}}$contains the factors uncorrelated both with the idiosyncratic residual and the custom factors.
We start by identifying what remains unexplained from the old factors F^X when we consider only the factors in F^Y. This is done through a simple linear regression model :

$$F^X = \tilde{\beta}^Y F^Y + \tilde{u} \tag{3.42}$$

The residuals $\tilde{u}$, that by construction are uncorrelated with F^Y, are defined as new factors $F^{\tilde{X}}$. These are the projections of the old factors F^X on the space orthogonal to F^Y:

$$F^{\tilde{X}} = F^X - \tilde{\beta}^Y F^Y = F^X[I - F^{'Y}(F^Y F^{'Y})^{-1} F^Y] \tag{3.43}$$

In this way we are able to decompose the return vector r in three uncorrelated quantities:

$$r = \begin{bmatrix} \beta^Y & \beta^{\tilde{X}} \end{bmatrix} \begin{bmatrix} F^Y \\ F^{\tilde{X}} \end{bmatrix} + \epsilon^i \tag{3.44}$$

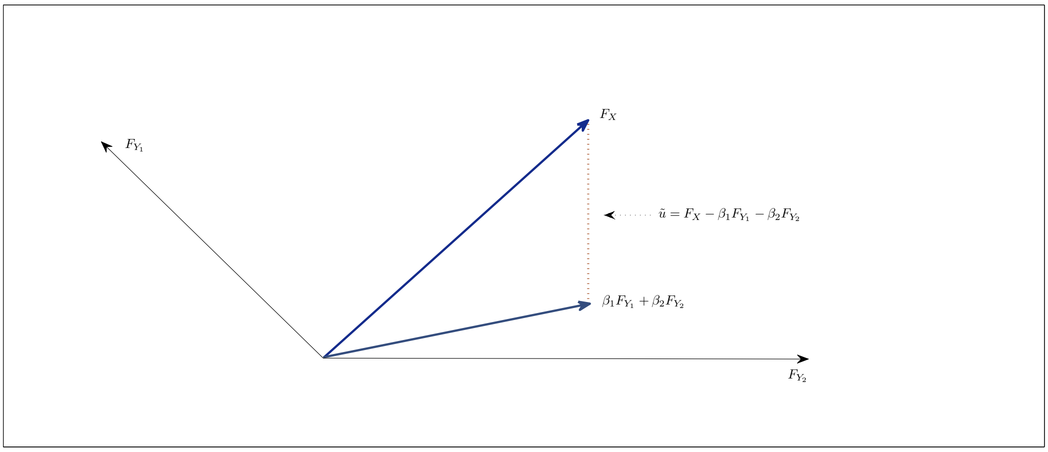


Figure 3.1: Geometric representation of the residuals $\tilde{u}$ orthogonal to the space generated by the custom factors F^Y

The estimated exposures are the result of an OLS regression of two uncorrelated groups of variables:

$$\begin{bmatrix} \beta^Y & \beta^{\tilde{X}} \end{bmatrix} = r \begin{bmatrix} F^Y \\ F^{\tilde{X}} \end{bmatrix}' \left(\begin{bmatrix} F^Y \\ F^{\tilde{X}} \end{bmatrix} \begin{bmatrix} F^Y & F^{\tilde{X}} \end{bmatrix} \right)^{-1} \tag{3.45}$$

The new introduced methodology is composed by three simple steps:
1) Given the custom factors F^Y and using equation (3.43) we obtain the new residual factors $F^{\tilde{X}}$.
2) The new exposures are :

$$\begin{bmatrix} \beta^Y & \beta^{\tilde{X}} \end{bmatrix} = \begin{bmatrix} rF^{Y'}(F^Y F^{Y'})^{-1} & rF^{\tilde{X}'}(F^{\tilde{X}} F^{\tilde{X}'})^{-1} \end{bmatrix} \tag{3.46}$$

3) As a final step, we compute the idiosyncratic residuals : $\epsilon^i = r - (\beta^Y F^Y + \beta^{\tilde{X}} F^{\tilde{X}})$. The procedure

is simple and doesn't require a lot of storage space. Once we have the return decomposition model, for the discussed risk measures the attribution to each factor is straightforward. Any portfolio manager can offer to his clients a set of custom factors and daily update the residual factors. In this way the client receives the information required and in the same time gets an idea of where does the residual risk come from. The new factors $F^{\tilde{X}}$ have a simple interpretation since they contain the returns that are not idiosyncratic and not explained from F^Y.

3.3.2 Application to the stock market

Here we apply the procedure described in the previous section to the equity market and in particular we consider the returns of the VFIAX fund index which replicates the performance of the SPX index. The evaluation date is 16/10/2012 and we use two years of historical data in computing the risk measures. As a first step we compute the exposures to the factors used for explaining the retuns, i.e daily log returns on GICS indexes.

Factor	COND	CONS	ENRS	FINL	HLTH	INDU	INFT	MATR	TELS	UTIL
β	0.110	0.130	0.122	0.141	0.098	0.112	0.187	0.040	0.024	0.040
	(0.007)	(0.009)	(0.005)	(0.004)	(0.008)	(0.008)	(0.006)	(0.006)	(0.005)	(0.006)

Table 3.3: Exposures of the VFIAX fund index to the 10 GICS sectors considered as factors computed considering the period 16/10/2010 till 16/10/2012.

The Global index return vector (quoted in the U.S market and downloaded from the Bloomberg terminal) is considered as a custom factor. We compute the projections of the GICS returns on the space orthogonal to that generated from the custom factor and applying the procedure introduced before, we get the idiosyncratic return vector. In figure 3.3.2 we plot the fund returns and the sequence of the idiosyncratic residuals. To confirm the fact,that to the residual term is attributed a small part of the fund risk we show the results of the decomposition of the Expected Shortfall in 3.3.2 for the three groups of factors : the custom, the projected and the idiosyncratic factors.

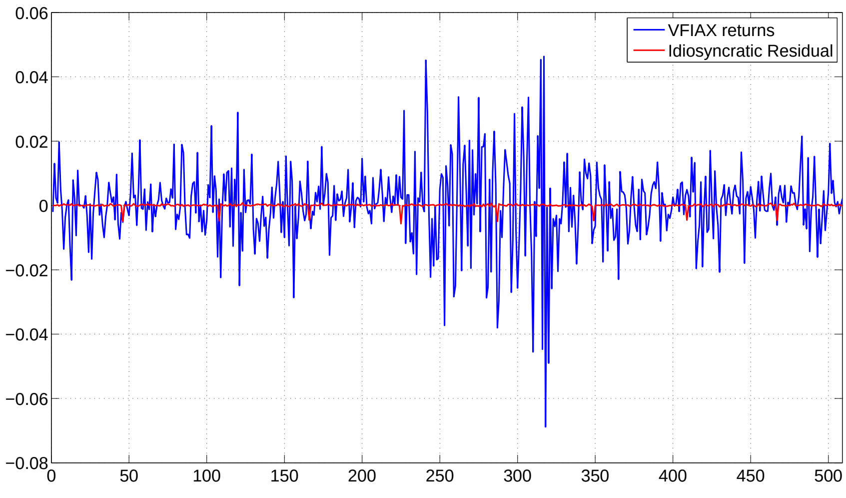


Figure 3.2: In this figure we show the time series of the VIAX returns and the idiosyncratic residual vector when the Global index is considered as custom factor.

	$VFIAX$	F^Y	$F^{\tilde{X}}$	ϵ^i
$ES_{0.05}$	0.0298	0.0249	0.0048	4.79E-05
		83.57%	16.27%	0.16%

Table 3.4: Decomposition of the VFIAX $ES_{0.05}$ considering as new factor F^Y the return on the Global index to which is most of the risk.

3.3.3 Application to fixed income securities

Here we give two examples of the methodology introduced in the previous section. In particular, we consider the returns of two bonds both with time to maturity 5 years paying coupons annually. The issuer of the first bond is assumed to have a z-spread similar to the Industrial Sector B rated companies while the second bond z-spread is that of the Utility sector A rated companies. Coherently, we assume a coupon rate of 6.875% for the first bond and a coupon rate of 2% for the second bond. The evaluation date is 16/11/2012 while the maturity date is 16/11/2017.
Given function g such that a bond price is :

$$P = g(\beta^X, F^X) \tag{3.47}$$

we can define $\beta^{X_i} = -MD^{X_i}$, i.e minus the modified duration of the i-th zero rate or i-th z-spread. The log-returns can be approximated using the modified duration associated to each risk factor :

$$r_t \approx -\sum_{i=1}^{5} MD^{z_{iy}} \Delta z_{iy} - MD^{z_{spread}} \Delta z_{spread} \tag{3.48}$$

The formula used to compute the modified duration is:

$$MD^{z_{iy}} = \frac{P(z_{iy} + 0.0001) - P(z_{iy})}{2P(z_{iy})0.0001} \tag{3.49}$$

We consider the daily realizations for the risk factors and use them to generate bond return scenarios. As custom factor we consider the returns on the Oil index whose prices are downloaded from the Bloomberg terminal.
We consider the Expected Shortfall as a risk measure. We expect the bond in the Industrial sector to have a higher exposure to the Oil index than the second bond. The two step-regression and the new projection factor method are used to determine the residuals. In the second approach, being the idiosyncratic residuals the part of return that can not be attributed neither to the custom factor nor to the projected factors, the risk contribution of the residual term should be lower.

At the evaluation date the price of the first bond is $P_1 = 96.92$ with z-spread = 645.39 basis points. The notional amount is N = 100 and the coupon rate = 6.875%. The original risk factors are the change in value of the zero rates corresponding to the payment maturities and of the z-spread. The new factor with respect to which we want to perform the return and risk decomposition is

	$\Delta\ z_{1y}$	$\Delta\ z_{2y}$	$\Delta\ z_{3y}$	$\Delta\ z_{4y}$	Δz_{5y}	$\Delta\ z_{spread1}$
Factor Exposure	0.067	0.125	0.174	0.216	3.864	4.447

Table 3.5: Here we report the factor exposures for the first bond that in this case are minus the modified duration for the considered zero rate or z-spread.

F^Y that is the log return of the Oil Index quoted in the market. We apply the two procedures discussed in the previous section, i.e the two-step regression and the orthogonal factor projection procedure. The residuals obtained with both methods are showed in figure 3.3 .

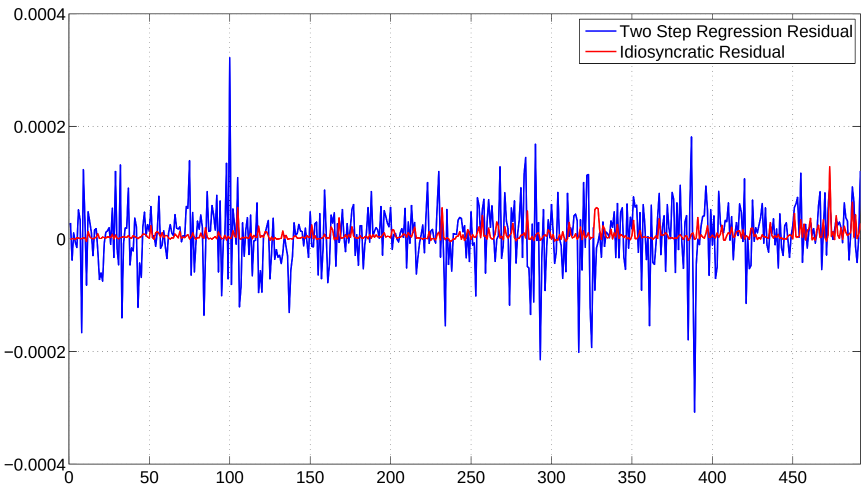


Figure 3.3: Comparison of the residual term in the two step regression and the idiosyncratic residual obtained with the second method for the bond with coupon rate 6.875%.The idiosyncratic errors are much more smaller than the residuals obtained in the two step regression method.

At the evaluation date the price of the second bond is $P_1 = 95.85$ with z-spread = 208.5 basis points. The notional amount is N = 100 and coupon rate = 2%. The original risk factors are the change in value of the zero rates corresponding to the payment maturities and of the z-spread. The new factor with respect to which we want to perform the return and risk decomposition is F^Y that is the log return of the Oil Index quoted in the market. We apply the two procedures discussed in the previous section, i.e the two-step regression and the orthogonal factor projection procedure. The residuals obtained with both methods are showed in fig 3.4.

	$BondIndu$	F^Y	$F^{\tilde{X}}$	ϵ^i
$ES_{0.05}(r)$	0.0128	1.70E-04	0.0127	-1.03E-04
		1.3%	99.5%	-0.8%

Table 3.6: Decomposition of the $ES_{0.05}$ of the bond in the Industrial sector considering as new factor F^Y the return on Oil index. The greater part of the risk is attributed to the projected factors $F^{\tilde{X}}$

.

	$\Delta\ z_{1y}$	$\Delta\ z_{2y}$	$\Delta\ z_{3y}$	$\Delta\ z_{4y}$	Δz_{5y}	$\Delta\ z_{spread2}$
Factor Exposure	0.021	0.040	0.059	0.076	4.675	4.871

Table 3.7: Here we report the factor exposures for the second bond that in this case are minus the modified duration for the considered zero rate or z-spread.

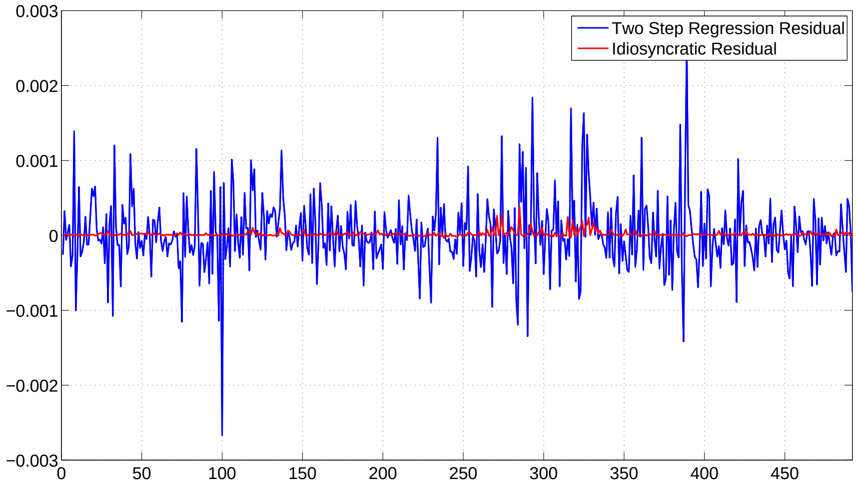


Figure 3.4: Comparison of the residual term in the two step regression and the idiosyncratic residual obtained with the second method for the bond with coupon rate 2%. The risk attributed to the custom factor here is smaller since the bond issuer operates in the Utility sector that is less influenced from the Oil index returns than the bond whose issuer operates in the Industrial sector.

	$BondIndu$	F^Y	$F^{\tilde{X}}$	ϵ^i
$ES_{0.05}(r)$	8.27E-03	1.80E-06	0.0083	-3.39E-05
		0.02%	100.39%	-0.41%

Table 3.8: Decomposition of the second bond $ES_{0.05}$ considering as new factor F^Y the return on Oil index. The greater part of the risk is attributed to the projected factors $F^{\tilde{X}}$

Chapter 4

Mixed Tempered Stable distribution

4.1 Semi-heavy tailed distributions

The Stable distribution has gained a great popularity in modeling economic and financial time series, starting from the seminal work of Mandelbrot [1963]. However, empirical evidence is not coherent with the Stable model assumption since the return distribution shows tails heavier than Normal but thinner than the Stable ones.

A drawback of the Stable distribution is that only fractional moments of order $p \leq \alpha$ with $\alpha \in (0, 2)$ exist and, consequently, the standard hypothesis for applying the Central Limit Theorem do not hold. For this reason several researchers have considered the Tempered Stable distribution as an valid alternative in modeling financial returns.
The Tempered Stable distribution can be obtained by multiplying the Lévy density of an α- Stable with a decreasing tempering function (see Cont and Tankov [2003]). Performing this operation, the tail behavior of the new distribution changes from heavy to semi-heavy characterized by exponential instead of power decay. The existence of the conventional moments is ensured and the Tempered Stable satisfies the conditions of the classical Central Limits Theorem. This is an advantage in modeling asset returns with respect to the Stable distribution. Indeed the return distributions tend to be closed to normal when we consider monthly or annual data (see Cont [2001] for survey of the Stylized Facts).

In this chapter, we propose a new distribution, namely the Mixed Tempered Stable (MixedTS henceforth). The idea is to build the new distribution in the similar way of the Normal Variance Mean Mixture (NVMM) where the Normal assumption is substituted by the standardized Tempered Stable. In this way, we are able to overcome some limits of the NVMM. In particular, the asymmetry and the kurtosis do not depend only on the mixing random variable but also on the

Tempered Stable.
Assuming that the mixing random variable follows a Gamma distribution, the proposed model has the Variance Gamma (see Madan and Seneta [1990b] and Loregian *et al.* [2012]), the Tempered Stable (see Cont and Tankov [2003]) and the Geo-Stable distributions (see Kozubowski *et al.* [1997]) as special cases. The new distribution is applied to real data. In particular, we conduct two experiments. In the first, we build a Garch(1,1) with MixedTS innovations and estimate it using the financial time series. In the second, we consider a multifactor model to describe the log returns of the Vanguard Fund Index which tries to replicate the performance of the S&P 500 index. As factors, we consider the Global Industry Classification Standard (GICS) indexes developed by MSCI-Barra provider. We capture the GICS dependence structure using the Independent Component Analysis introduced by Comon [1994] and developed by Hyvarinen *et al.* [2001].
Once fixed the number of the independent components, we use the Mixed Tempered Stable to model each of them. Since the observed fund returns are modeled as a linear sum of the independent ones, the single factor return density could be non-gaussian and/or semi-heavy tailed. This is confirmed by our results and the different values obtained for the MixedTS parameters suggest that the component distributions can be of quite different nature. The Independent Component Analysis has been already used in finance to model the interest rates term structure (see Bellini and Salinelli [2003]). Recently Madan [2006b] used a non-Gaussian factor model and modelled the components, coming fromman ICA analysis, using the Variance Gamma distribution. Our multifactor model can be seen as a generalization of this latter since, as observed above, the Variance Gamma is a particular case of the Mixed Tempered Stable distribution. This portfolio return decomposition can be used in risk management. Marginal contribution of each linear transformed factor can be easily computed for given homogeneous risk measures as described in Chapter 3. The outline of the chapter is as follows. In Section 2 we describe the main characteristics of the Tempered Stable distribution and make some observations useful for the derivation of the new distribution. In Section 3 we introduce the Mixed Tempered Stable distribution, its main features and illustrate the fitting of the MixedTS distribution to the Independent Components. Section 5 gives some results for the computation of the risk measures considered in this thesis considering the new distribution.

4.2 Tempered Stable distribution

In this section we review the main features of the Tempered Stable distribution. A random variable X follows a Tempered Stable distribution if its Lévy measure is given by:

$$\nu\left(dx\right) = \left(\frac{C_{+}e^{-\lambda_{+}x}}{x^{1+\alpha_{+}}}1_{x>0} + \frac{C_{-}e^{-\lambda_{-}|x|}}{|x|^{1+\alpha_{-}}}1_{x<0}\right)dx \tag{4.1}$$

with $\alpha_+,\ \alpha_- \in (0,2)$ and $C_+,\ C_-,\ \lambda_+,\ \lambda_- \in (0,+\infty)$.
The characteristic function is obtained by solving the integral [see Cont and Tankov, 2003]:

$$\begin{aligned} E\left[e^{iuX}\right] &= \exp\left[iu\gamma + \int_{\Re}\left(e^{iux} - 1 - iux\right)\nu\,(dx)\right] \\ &= \exp\left\{iu\gamma + C_+\Gamma\,(-\alpha_+)\left[(\lambda_+ - iu)^{\alpha_+} - \lambda^{\alpha_+}\right]\right. \\ &+ \left.C_-\Gamma\,(-\alpha_-)\left[(\lambda_- + iu)^{\alpha_-} - \lambda^{\alpha_-}\right]\right\} \end{aligned} \tag{4.2}$$

where $\gamma \in \Re$. As observed in Küchler and Tappe [2013], for $\alpha_+,\ \alpha_- \in (0,1)$, the Tempered Stable is obtained as a difference of two independent *one sided* tempered stable distributions introduced in Tweedie [1984]. The corresponding process shows finite variation with infinite activity.

The interest of researchers for the Tempered Stable distribution is confirmed by the fact that the following cases have been investigated in literature:

- For $C_+ = C_- = C$ and $\alpha_+ = \alpha_- = \alpha$, the CGMY distribution is obtained [see Carr *et al.*, 2002].
- If $\alpha_+ = \alpha_-$ and $\lambda_+ = \lambda_-$ we get the truncated Lévy flight introduced in Koponen [1995].
- Computing the limit for $\alpha_+ = \alpha_- \to 0^+$ we get the Bilateral Gamma distribution [see Küchler and Tappe, 2008a,b, 2009].
- Taking the limit for $\alpha_+ = \alpha_- \to 0^+$, $C_+ = C_-$ and $\lambda_+ = \lambda_-$ the Variance Gamma is obtained [see Madan and Seneta, 1990a; Loregian *et al.*, 2012, for estimation].

In this chapter, we consider the same restrictions as in [see Kim *et al.*, 2008], i.e : $\alpha_+ = \alpha_- = \alpha$ and $\gamma = \mu - \Gamma\,(1-\alpha)\left(C_+\lambda_+^{\alpha-1} - C_-\lambda_-^{\alpha-1}\right)$. In this case the distribution of $X \sim CTS\,(\alpha,\ \lambda_+,\ \lambda_-,\ C_+,\ C_-,\ \mu)$ is called classical tempered stable and $E\,(X) = \mu$. Its characteristic function is given by:

$$\begin{aligned} E\left[e^{iuX}\right] &= \Phi\,(u;\ \alpha,\ \lambda_+,\ \lambda_-,\ C_+,\ C_-,\ \mu) \\ &= \exp\left[iu\mu - iu\Gamma\,(1-\alpha)\left(C_+{\lambda_+}^{\alpha-1} - C_-{\lambda_+}^{\alpha-1}\right)\right. \\ &+ \left.C_+\Gamma\,(-\alpha)\left((\lambda_+ - iu)^{\alpha} - \lambda_+^{\alpha}\right) + C_-\Gamma\,(-\alpha)\left((\lambda_- + iu)^{\alpha} - \lambda_-^{\alpha}\right)\right] \end{aligned}$$

The cumulants of the r.v X can though be obtained taking the derivatives of the characteristic function :

$$c_n\,(X) := \frac{1}{i^n}\frac{\partial^n}{\partial u^n}\ln\left(E\left[e^{iuX}\right]\right)\Bigg|_{u=0} \tag{4.3}$$

Given the distribution parameters, for $n = 1$ we have :

$$c_1(X) = \mu \tag{4.4}$$

and for $n \geq 2$:

$$c_n(X) = \Gamma(n-\alpha)(C_+\lambda_+^{\alpha-n} + (-1)^n C_-\lambda_-^{\alpha-n}) \tag{4.5}$$

Therefore using (4.5) the following quantities are computed:

$$\begin{cases} E\left(X\right) = c_1\left(X\right) = \mu \\ Var\left(X\right) = c_2\left(X\right) = \Gamma\left(2-\alpha\right)\left[C_+\lambda_+^{\alpha-n} + \left(-1\right)^n C_-\lambda_-^{\alpha-n}\right] \\ \gamma_1 = \frac{c_3(X)}{c_2^{3/2}(X)} = \frac{\Gamma(3-\alpha)\left[C_+\lambda_+^{\alpha-3} - C_-\lambda_-^{\alpha-3}\right]}{c_2^{3/2}(X)} \\ \gamma_2 = 3 + \frac{c_4(X)}{c_2^2(X)} = 3 + \frac{\Gamma(4-\alpha)\left[C_+\lambda_+^{\alpha-4} - C_-\lambda_-^{\alpha-4}\right]}{c_2^2(X)} \end{cases} \tag{4.6}$$

From the skewness formula it can be noticed that the difference between $C_+\lambda_+^{\alpha-3}$ and $C_-\lambda_-^{\alpha-3}$ drives the asymmetry while for $C_+ = C_-$we must look only at the two tempering parameters λ_+ and λ_-.

Remark 9. *Fixing $\lambda_+ = \lambda_- = \lambda$, $\mu = 0$ and $C_+ = C_- = C$, the characteristic function of a symmetric Tempered Stable distribution is given by:*

$$E\left[\exp\left(iuX\right)\right] = \exp\left[iu\mu + C\Gamma\left(1-\alpha\right)\left[\left(\lambda - iu\right)^\alpha + \left(\lambda + iu\right)^\alpha - 2\lambda^\alpha\right]\right] \tag{4.7}$$

If $r = \sqrt{u^2+\lambda^2}$ and $\theta = arctg\left(\frac{u}{\lambda}\right)$, the characteristic function in (4.7) *becomes:*

$$\begin{aligned} E\left[\exp\left(iuX\right)\right] &= \exp\left[C\Gamma\left(1-\alpha\right)\left[r^\alpha e^{-i\alpha\theta} + r^\alpha e^{i\alpha\theta} - 2\lambda^\alpha\right]\right] \\ &= \exp\left[C\Gamma\left(1-\alpha\right)\left[2r^\alpha\cos(\alpha\theta) - 2\lambda^\alpha\right]\right] \end{aligned}$$

The last equality holds due to the Euler relation $\frac{e^{i\theta}+e^{-i\theta}}{2} = \cos\left(\alpha\theta\right)$. Taking the limit for $\lambda \to 0^+$ we obtain the symmetric $\alpha-$stable distribution since:

$$\lim_{\lambda\to 0^+}\exp\left[C\Gamma\left(1-\alpha\right)\left[2\left(u^2+\lambda^2\right)^{\frac{\alpha}{2}}\cos(\alpha\theta) - 2\lambda^\alpha\right]\right] = \exp\left[C\Gamma\left(1-\alpha\right)\left[2\left|u\right|^\alpha\cos\left(\frac{\alpha\pi}{2}\right)\right]\right]$$

The r.v X has zero mean and unit variance for $\mu = 0$ and :

$$C = C_+ = C_- = \frac{1}{\Gamma(2-\alpha)(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2})} \tag{4.8}$$

The correponding r.v is standardized, i.e $X \sim stdCTS(\alpha, \lambda_+, \lambda_-)$. It useful to observe that in the standardized CTS distribution C is fully determined once given the values for α, λ_+ and λ_-. Its characteristic exponent, defined as $L_{stdCTS}\left(u;\alpha,\lambda_+,\lambda_-\right) = logE[e^{iuX}]$, is given by:

$$L_{stdCTS}(u:\alpha,\lambda_+,\lambda_-) = \frac{\left(\lambda_+ - iu\right)^\alpha - \lambda_+^\alpha + \left(\lambda_- + iu\right)^\alpha - \lambda_-^\alpha}{\alpha(\alpha-1)(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2})} + \frac{iu(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1})}{(\alpha-1)(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2})} \tag{4.9}$$

For $\alpha \to 2$ the limiting distribution coincides with the normal distribution:

$$\lim_{\alpha\to 2} L_{stdCTS}(u:\alpha,\lambda_+,\lambda_-) = -\frac{u^2}{2}$$

Remark 10. *Condition* (4.8) *implies that the convergence to* $\alpha-$ *stable distribution is not possible since the characteristic exponent would converge to zero.*

The standardized Classical Tempered Stable distribution has the following property that makes it appealing for mixtures.

Proposition 11. *Let* $\tilde{X} \sim stdCTS\left(u : \alpha, \lambda_+\sqrt{h}, \lambda_-\sqrt{h}\right)$ *and* $h \in (0, +\infty)$ *then the random variable* $Y \stackrel{d}{=} \sqrt{h}\tilde{X}$ *has the following characteristic exponent:*

$$\ln E\left[e^{iuY}\right] = h\left[\frac{(\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - \lambda_-^\alpha}{\alpha(\alpha-1)\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)} + \frac{iu\left(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}\right)}{\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)}\right] \tag{4.10}$$

Moreover if $h \in \mathbb{N}$ *we have:*

$$Y \stackrel{d}{=} \sum_{j=1}^{h} X_j \tag{4.11}$$

where X_j *are iid* $stdCTS\left(\alpha,\ \lambda_+,\ \lambda_-\right)$

Proof. To derive the (4.10) exponent we evaluate (4.9) in $\sqrt{h}u$ and obtain:

$$\ln E\left[e^{iuY}\right] = \left[\frac{\left(\lambda_+\sqrt{h} - iu\sqrt{h}\right)^\alpha - h^{\frac{\alpha}{2}}\lambda_+^\alpha + \left(\lambda_-\sqrt{h} + iu\sqrt{h}\right)^\alpha - h^{\frac{\alpha}{2}}\lambda_-^\alpha}{\alpha(\alpha-1)\left(h^{\frac{\alpha}{2}-1}\lambda_+^{\alpha-2} + h^{\frac{\alpha}{2}-1}\lambda_-^{\alpha-2}\right)} + \frac{i\sqrt{h}u\left(h^{\frac{\alpha-1}{2}}\lambda_+^{\alpha-1} - h^{\frac{\alpha-1}{2}}\lambda_-^{\alpha-1}\right)}{\left(h^{\frac{\alpha}{2}-1}\lambda_+^{\alpha-2} + h^{\frac{\alpha}{2}-1}\lambda_-^{\alpha-2}\right)}\right] \tag{4.12}$$

factorizing h we get:

$$\ln E\left[e^{iuY}\right] = \left[\frac{h^{\frac{\alpha}{2}}\left(\lambda_+ - iu\right)^\alpha - \lambda_+^\alpha + (\lambda_- + iu)^\alpha - \lambda_-^\alpha}{h^{\frac{\alpha}{2}-1}\alpha(\alpha-1)\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)} + \frac{iuh^{\frac{\alpha}{2}}\left(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}\right)}{h^{\frac{\alpha}{2}-1}\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)}\right] \tag{4.13}$$

simplifying we obtain the result (4.10).
To prove the (4.11) we use the iid assumption for X_j and the characteristic exponent of the random variable $\sum_{j=1}^{h} X_j$ becomes:

$$\ln E\left[\exp\left(\sum_{j=1}^{h} X_j\right)\right] = h \ln E\left[\exp\left(X_1\right)\right] \tag{4.14}$$

where $X_1 \sim stdCTS\left(\alpha,\ \lambda_+,\ \lambda_-\right)$. Using (4.9) we obtain the characteristic exponent of $\sqrt{h}\tilde{X}$ that implies the condition (4.11). □

4.3 Mixed Tempered Stable distribution

In this section, using the prop. 11, we build a new distribution that is shown to have some nice mathematical and statistical chacteristics .

4.3.1 Definition and particular cases

Definition 12. *We say that a continuous random variable* **Y** *follows a Mixed Tempered Stable distribution if:*

$$Y \stackrel{d}{=} \sqrt{V}\tilde{X} \tag{4.15}$$

where $\tilde{X}\,|V \sim stdTS(\alpha, \lambda_+\sqrt{V}, \lambda_-\sqrt{V})$. *V is a Lévy distribution defined on positive axis and its m.g.f always exists.*

The logarithm of the m.g.f. is :

$$\Phi_V(u) = \ln\left[E\left[\exp\left(uV\right)\right]\right] \tag{4.16}$$

We compute the characteristic exponent for the new distribution and apply the law of iterated expectation:

$$\begin{aligned} E\left[e^{iu\sqrt{V}\tilde{X}}\right] &= E\left\{E\left[e^{iu\sqrt{V}\tilde{X}}\middle| V\right]\right\} \\ &= \exp\left[\Phi_V\left(L_{stdCTS}\left(u;\ \alpha,\ \lambda_+,\ \lambda_-\right)\right)\right] \end{aligned} \tag{4.17}$$

The characteristic function identifies the distribution at time one of a time changed Lévy process.From [see Sato, 1998; Carr and Wu, 2004] the distribution is infinitely divisible.

Proposition 13. *The first four moments for the MixedTS are:*

$$\begin{cases} E\left[\sqrt{V}\tilde{X}\right] = 0 \\ Var\left[\sqrt{V}\tilde{X}\right] = E\left[V\right] \\ \gamma_1 = \left(2-\alpha\right)\frac{\left(\lambda_+^{\alpha-3}-\lambda_-^{\alpha-3}\right)}{\left(\lambda_+^{\alpha-2}+\lambda_-^{\alpha-2}\right)}E^{-1/2}\left[V\right] \\ \gamma_2 = \left[3+\left(3-\alpha\right)\left(2-\alpha\right)\frac{\left(\lambda_+^{\alpha-4}+\lambda_-^{\alpha-4}\right)}{\left(\lambda_+^{\alpha-2}+\lambda_-^{\alpha-2}\right)}\right]\frac{E\left[V^2\right]}{E^2\left[V\right]} \end{cases}$$

Figure 4.1 shows the behaviour of the skewness for different combinations of λ_+ and λ_- and fixed α.

In figure 4.2 we have the behaviour of the kurtosis for different values of $\lambda_+,\ \lambda_-$ and α.

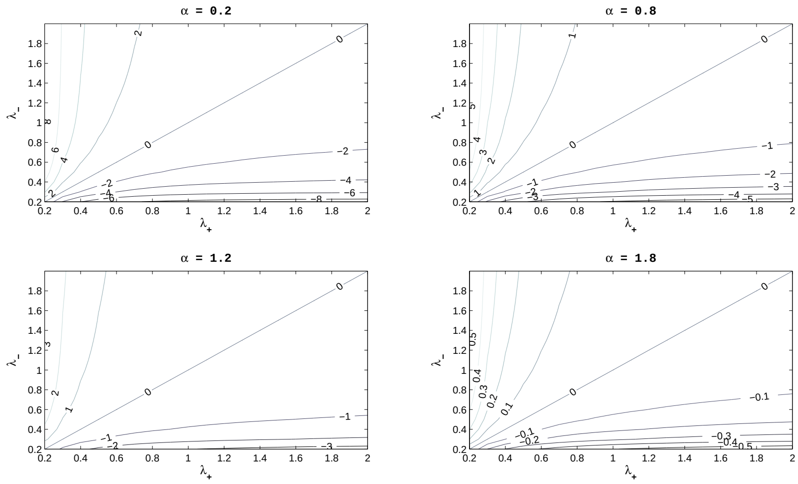


Figure 4.1: Consider the case when $V \sim \Gamma(1,1)$ and fix some value for α. We plot the skewness curve levels for different combinations of λ_+ and λ_- to have an idea of the possible skewness values. In the particular case when they coincide, the skewness is zero. The effect of an higher α is the reduction of the skewness level (kept fixed values of the other parameters). The distribution of the MixedTS becomes symmetric for $\alpha = 2$.

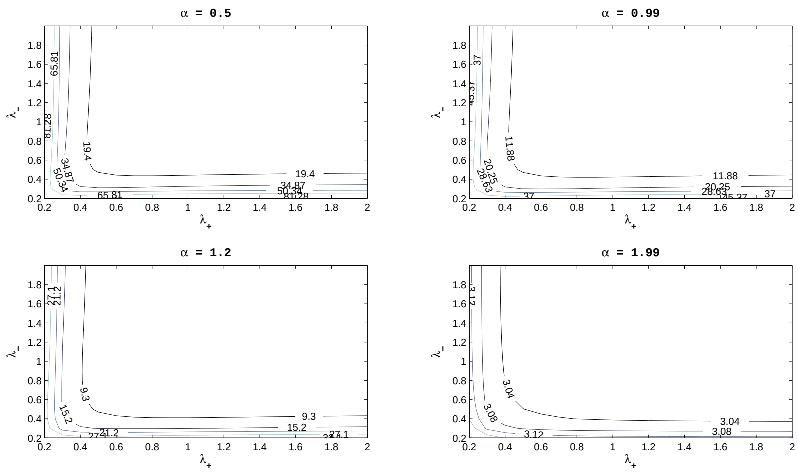


Figure 4.2: Consider the case when $V \sim \Gamma(1,1)$ and fix some value for α. We plot the kurtosis values for different combinations of λ_+ and λ_- to have an idea of the possible kurtosis values. The effect of an higher α is the reduction of the kurtosis level. If the fixed value for α is 1.9 the curve level for kurtosis tend to be close to 3 and the limiting case of *kurtosis*=3 is obtained for $\alpha = 2$.

If we assume that $V \sim \Gamma(a, \sigma^2)$ the characteristic exponent in (4.17) becomes:

$$\begin{aligned} E\left[\exp\left(u\sqrt{V}X\right)\right] &= \exp\left[-a\ln\left(1-\sigma^2\frac{(\lambda_+ - iu)^\alpha - (\lambda_+)^\alpha + (\lambda_- + iu)^\alpha - (\lambda_-)^\alpha}{\alpha\left(\alpha-1\right)\left((\lambda_+)^{\alpha-2}+(\lambda_-)^{\alpha-2}\right)}\right.\right. \\ &- \left.\left.\sigma^2\frac{iu(\lambda_+^{\alpha-1}-\lambda_-^{\alpha-1})}{(\alpha-1)(\lambda_+^{\alpha-2}+\lambda_-^{\alpha-2})}\right)\right] \end{aligned} \tag{4.18}$$

For $\sigma = \frac{1}{\sqrt{a}}$ we compute the limit for a going to infinity and obtain the *stdCTS* as a special case (see Figure 4.3(b)). The symmetric Variance Gamma distribution is obtained by choosing $\alpha = 2$ as shown in Figure 4.3(a).

Proposition 14. *Choosing*

$$\begin{aligned} \lambda_+ &= \lambda_- = \lambda \\ a &= 1 \qquad\qquad (4.19) \\ \sigma &= \lambda^{\frac{\alpha-2}{2}}\gamma^{\frac{\alpha}{2}}\sqrt{\left|\frac{\alpha\left(\alpha-1\right)}{\cos\left(\alpha\frac{\pi}{2}\right)}\right|} \end{aligned} \tag{4.20}$$

and computing the limit for $\lambda \to 0^+$ we obtain the Geometric Stable distribution.

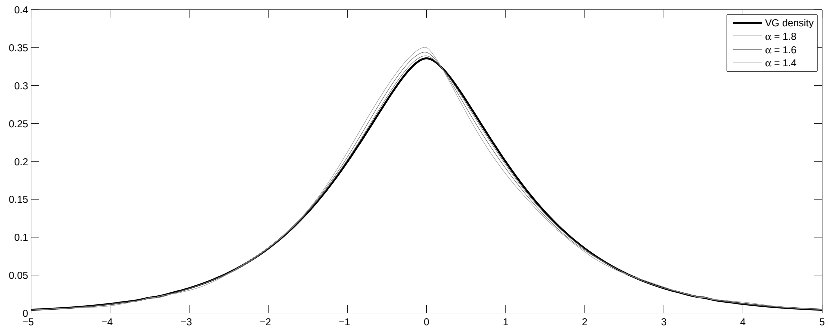
Indeed substituting the conditions (4.20) in the characteristic exponent of the new distribution we get:

$$E\left[\exp\left(iu\sqrt{V}X\right)\right] = \exp\left[-\ln\left(1-\frac{(\lambda_+ - iu)^\alpha - (\lambda_+)^\alpha + (\lambda_- + iu)^\alpha - (\lambda_-)^\alpha}{2\alpha\left(\alpha-1\right)}\right)\right]$$

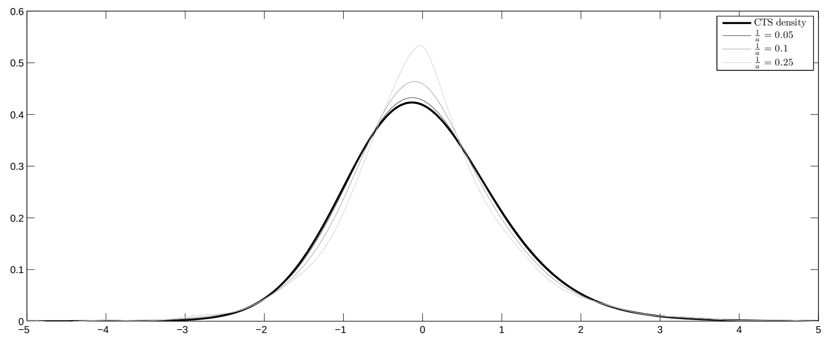
Applying the limit and following the same arguments for the convergence of the Tempered Stable to the symmetric distribution we get:

$$E\left[\exp\left(iu\sqrt{V}X\right)\right] \to \left(1-|u|^\alpha\frac{cos\left(\alpha\frac{\pi}{2}\right)}{\alpha\left(\alpha-1\right)}\right)^{-1}$$

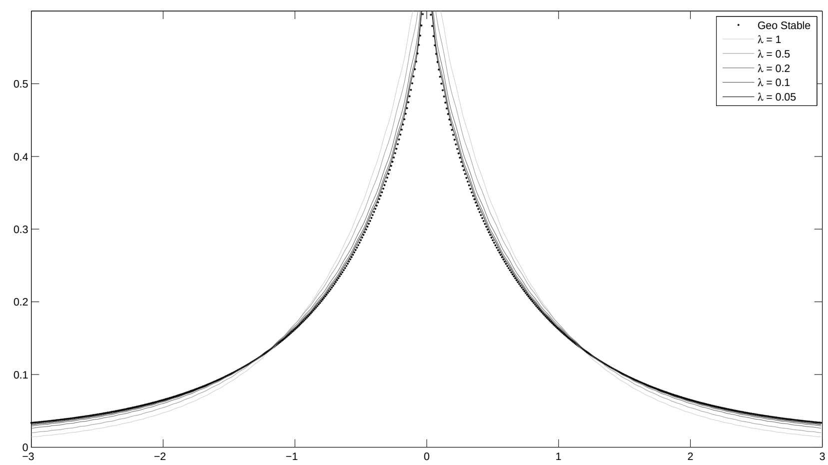
for $\alpha \neq 1$. The convergence to the Geometric Stable distribution is shown in Figure 4.3(c).



(a) The symmetric VG distribution is a particular case of the MixedTS and it is obtained for $\alpha = 2$. We fix$\mu_0 = 0$, $\mu = 0$, $\sigma = 1.2$, $a = 1.7$, $\lambda_+ = 1.2$ and $\lambda_- = 8$. In the figure we plot the MixedTS for different α values. The distribution associated is asymmetric but the limiting case not.



(b) The symmetric CTS distribution is a particular case of the MixedTS and it is obtained for $\alpha = 2$. We fix$\mu_0 = 0$, $\mu = 0$, $\lambda_+ = 1.2$ and $\lambda_- = 8$ and $\alpha = 1.4$. In the figure we plot the MixedTS for different values of a.



(c) To show the convergence of the MixedTS distribution to the Geometric Stable one we plot it for different λ values given $\mu_0 = 0$, $\mu = 0$, $a = 1$ and $\gamma = 1$. Recall that the GeoStable is obtained as a limiting case when , $\sigma = \lambda^{\frac{\alpha-2}{2}} \gamma^{\frac{\alpha}{2}} \sqrt{\left| \frac{\alpha(\alpha-1)}{\cos(\alpha \frac{\pi}{2})} \right|}$ and $\lambda \to 0$. Observe that as λ gets smaller the tails get heavier. We cut the plot since in zero the GeoStable distribution has a peak going to $+\infty$.

Figure 4.3: Special cases of the Mixed Tempered Stable distribution

4.3.2 Investigation using real market data

This section is devoted to the empirical investigation of the MixedTS distribution in modeling asset returns. We consider daily log returns of the Vanguard Fund Index which tries to replicate the Standard and Poor 500 performance. It seems quite natural to consider as portfolio risk factors the daily log-returns of the 10 GICS sector indexes since each of the S&P500 members belongs to one of them. The data are daily log returns ranging from 14-June-2010 to 20-September-2012 obtained from the Bloomberg data provider. We denote by $X^{(t)}$ the vector of risk factors at time t with the i-th component $X_i^{(t)} = logS_i^{(t)}/S_i^{(t-1)}$ being the i-th GICS daily log-return. We can now define a linear model for the portfolio returns:

$$r_P = \beta X + \epsilon. \tag{4.21}$$

where β is the $1 \times n$ vector of factor exposures, X is the $n \times t$ matrix of factor returns and ϵ is the $1 \times t$ vector of residuals
Note that we drop the t index since we report in matrix notations the whole time-series for our quantities. In the simplest case β and ϵ are obtained through an OLS regression. Here the number of risk factors is $n = 10$ and the residuals are centered due to the particular choice of our portfolio. The R^2 is about 99 per cent meaning that the explained variance is quite high.In Table 4.1 we report factor exposures and observe that they are coherent with the market capitalization associated to each sector. We apply the FastICA Algorithm to the GICS returns and find the components that

Regression coefficients and Capitalization weights										
	COND	**CONS**	**ENRS**	**FINL**	**HLTH**	**INDU**	**INFT**	**MATR**	**TELS**	**UTIL**
β	0.1105	0.1154	0.1238	0.1442	0.105	0.1145	0.1818	0.0378	0.0220	0.0415
Cap weight (14/06/2010)	0.1103	0.1165	0.1206	0.1441	0.1190	0.1221	0.1800	0.0115	0.0213	0.0546
Cap weight (20/09/2012)	0.1108	0.1089	0.1127	0.1507	0.1228	0.099	0.1921	0.03499	0.03175	0.0362

Table 4.1: We perform a regression analysis and obtain the factor exposures for our portfolio. The dataset is composed by closing prices ranging from 14/06/2010 to 21/09/2012. The associated R^2 is $99,69$ meaning that the explanatory power of our factors is quite high. We report the capitalization weight at the begining and at the end of the study period. The factor exposures are in line with the average market capitalization for each sector.

will be the ones that maximize the non-gaussianity condition present in the optimization algorithm. If we think the ICs to be the columns of a matrix S, the risk factors time series can be seen as linear transformations of the independent signal sources. In the mixing matrix $A \in \Re^{n \times n}$ is contained the information about the weight of the single original source in the market sector, i.e $X = AS$.

The portfolio return distribution can be obtained as a linear combination of n independent distributions. The Central Limit theorem suggests that the distribution of the single IC can be heavy tailed and/or asymmetric since what we observe is the realization vector of a sum of distributions.
We use the Jarque-Bera test to check for non-normality in the univariate case. Simultaneously

sample skewness and kurtosis are compared with the corresponding normal values. In Table 4.2 we show the sample skewness, kurtosis and p-values for the Jarque-Bera test of the 10 ICs.

	Mean	Std	Skew	Ex-Kurt	Min	Max
VFIAX	0.0005	0.0121	-0.4693	3.9683	-0.0688	0.0463
COND	0.0007	0.0129	-0.5674	3.0566	-0.0690	0.0472
CONS	0.0006	0.0078	-0.3871	3.2413	-0.0390	0.0332
ENRS	0.0006	0.0158	-0.3878	3.3909	-0.0864	0.0687
FINL	0.0002	0.0175	-0.3415	4.1493	-0.1052	0.0789
HLTH	0.0006	0.0101	-0.4205	3.9130	-0.0540	0.0456
INDU	0,0004	0.0142	-0.4390	2.7993	-0.0711	0.0495
INFT	0.0007	0.0130	-0.2948	2.0849	-0.0596	0.0445
MATR	0.0005	0.0159	-0.3575	2.5824	-0.0756	0.0593
TELS	0.0007	0.0097	-0.2897	2.9581	-0.0550	0.0426
UTIL	0.0004	0.0089	-0.1124	4.7863	-0.0563	0.0414

Table 4.2: The reported statistics for the fund VFIAX and for the GICS show that the corresponding distributions are negatively skewed and with tails heavier than the one generated from a normal distribution.

The portfolio return can be decomposed in the form:

$$r_P = \beta^F F + \beta^N N + \epsilon. \tag{4.22}$$

with $F \in \Re^{l \times t}$ being the matrix containing the l rows of the S matrix containing the components we decided to be meaningful in the market and with $N \in \Re^{(n-l)\times t}$ the remaining considered as noise. The new exposures β^F and β^N are obtained by multiplying of the initial exposures β with the associated rows in the mixing matrix A.
The linear decomposition of returns in terms of uncorrelated factors eases the computation of the characteristic function:

$$E[e^{iur_p}] = E[e^{iu[\sum_{i=1}^{l} \beta_i^F F_i + \hat{\epsilon}}]] = \prod_{i=1}^{l} E[e^{iu\beta_i^F F_i}]E[e^{iu\hat{\epsilon}}] \tag{4.23}$$

The linear relation in the portfolio return model can be used to compute the marginal contribution to return/risk of each of the chosen IC or in a portfolio optimization problem as in Madan [2006a]. We emphasize the fact that our main focus is not introducing a new method on how to use ICA in finance but to stress how the flexibility of the Mixed Tempered Stable distribution can allow to capture contemporaneously the different shapes of each IC.

As measures of fit we consider the Mortara index A_1 and the quadratic K. Pearson index A_2 :

$$A_1 = \frac{1}{n}\sum_{j=1}^{S}|n_j - \hat{n}_j|$$

$$A_2 = \sqrt{\frac{1}{n}\sum_{j=1}^{S}\frac{(n_j - \hat{n}_j)^2}{\hat{n}_j}}$$

where S is the number of classes, n_j are the observed frequencies while $\hat{n}_j$ are the theoretical ones. The n is the summation of the theoretical frequencies.

Table 4.3 reports the estimated mixing matrix obtained using the FastICA algorithm.

Mixing Matrix									
I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**
-0.0035	-0.0098	-0.0016	0.0024	-0.0018	0.0035	-0.0044	-0.0014	0.0028	0.0023
-0.0008	-0.0059	-0.0001	0.0030	-0.0007	0.0029	0.0002	-0.0006	0.0021	-0.0014
0.0030	-0.0126	-0.0006	0.0013	0.0000	0.0029	-0.0067	-0.0034	0.0034	-0.0015
-0.0022	-0.0149	-0.0020	0.0047	0.0023	0.0030	-0.0052	-0.0027	-0.0027	0.0005
-0.0021	-0.0083	-0.0017	0.0029	-0.0005	0.0002	-0.0026	0.0005	0.0027	-0.0020
-0.0036	-0.0103	-0.0016	0.0029	0.0018	0.0035	-0.0055	-0.0037	0.0031	-0.0009
-0.0028	-0.0089	-0.0030	0.0027	-0.0038	0.0019	-0.0046	-0.0050	0.0015	-0.0008
-0.0019	-0.0113	-0.0023	0.0005	-0.0003	0.0060	-0.0082	-0.0018	0.0013	-0.0028
-0.0029	-0.0077	0.0044	0.0013	-0.0010	0.0006	-0.0012	-0.0011	0.0007	-0.0017
-0.0012	-0.0080	-0.0009	-0.0008	0.0004	0.0014	0.0018	-0.0011	0.0018	-0.0011

Table 4.3: Estimated Mixing Matrix obtained applying the FastIca algorithm to the dataset composed by the GICS. Only 8 out of the 10 Independent Components are chosen to use as factors. The remaining two are considered as noise.

The empirical densities of the independent components are shown in figure 4.4. In table 4.5 we give the MixedTS fitted parameters for each component and some measures of fit.

Statistics Ics										
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**	**IX**	**X**
Skewness	-0,6496	-0,1200	-0,5531	0,2913	-0,0349	-0,2916	0,0876	-0,1975	-0,0881	-0,0021
Kurtosis	7,7030	7,5633	5,9752	5,9352	4,6628	4,1283	4,2410	3,7250	3,6370	3,4420
JB-Statistic	546,5730	479,4040	231,3230	205,6020	63,5910	37,0450	36,0660	15,6540	10,0440	4,4830

Table 4.4: We report the skewness, kurtosis and JB-test for each component.

The four ICs with the highest JB statistic are considered as factors while the remaining ones as noise. The VFIAX return density was reconstructed using the MixedTS distribution for the factors

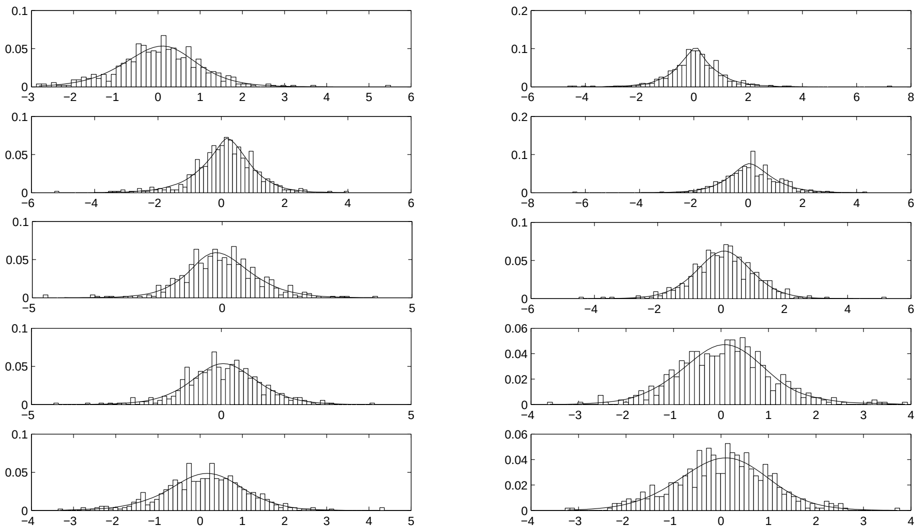


Figure 4.4: The MixedTS was fitted to each IC empirical density. The fitted parameters are reported in 4.5.It is easy to observe that the new distribution contemporaneously captures density shapes which are asymmetric and/or fat tailed.

and assuming normality for the noise. For comparison we plot the normal distribution fitted to the fund return density in figure 4.5.

Mixed TS Parameters and Fitting Measures										
	I	II	III	IV	V	VI	VII	VIII	IX	X
σ	1.0146	0.6666	0.7919	0.7814	0.8071	0.5983	0.6371	0.2595	0.1907	0.6358
a	1.2686	2.3377	2.1517	2.0305	1.4409	3.2953	2.8364	14.7360	30.4567	2.5446
λ_+	1.0247	0.9146	1.8861	1.0256	1.0000	1.8062	2.2127	8.2816	9.9490	1.0890
λ_-	0.9965	1.3665	0.1000	1.0431	1.0000	0.1000	0.6080	0.1000	0.1000	1.2074
α	1.3724	1.7579	1.7150	1.5458	1.3000	1.6660	1.5057	1.7187	0.5000	1.8469
A_2	0.0082	0.0064	0.0059	0.0072	0.0076	0.0036	0.0077	0.0062	0.0052	0.0063
X_2	0.0463	0.0781	0.0471	0.0512	0.0439	0.0314	0.0482	0.0434	0.0314	0.0412
A_1	0.0048	0.0040	0.0038	0.0052	0.0045	0.0026	0.0047	0.0043	0.0037	0.0038

Table 4.5: We fitted the MixedTS distribution to the ICs considered as factors in our model. In the table we show the parameters and the fitting measures.

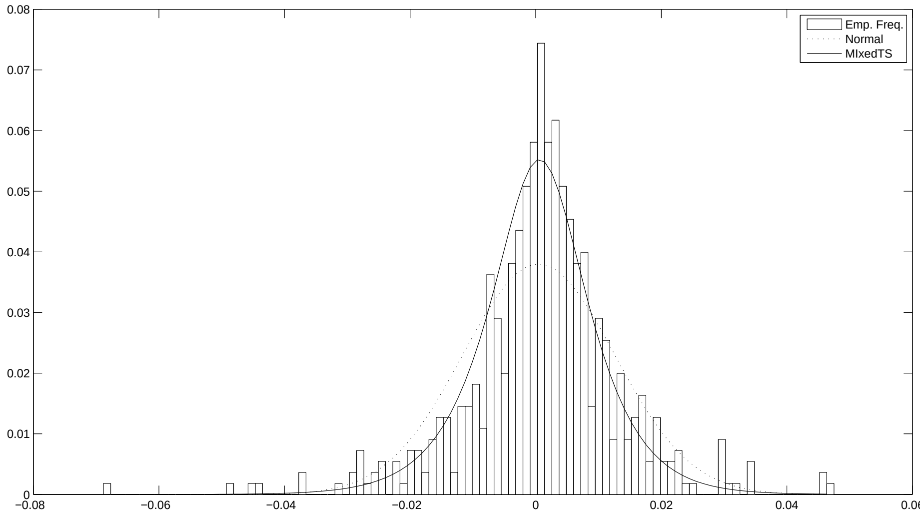


Figure 4.5: The four ICs with the highest JB statistic are considered as factors while the remaining ones as noise. The VFIAX return density was reconstructed using the MixedTS distribution for the factors and assuming normality for the noise. For comparison we plot the normal distribution fitted to the fund return density.

	VFIAX	**COND**	**CONS**	**ENRS**	**FINL**	**HLTH**	**INDU**	**INFT**	**MATR**	**TELS**	**UTIL**
μ_0	-0.0681	-0.0318	0.1399	-0.3049	-0.0830	-0.0338	-0.0655	0.0064	-0.0212	0.6256	0.0936
μ	0.0601	0.0227	-0.0454	0.1310	0.0204	0.0780	0.0232	-0.0311	0.0605	-0.1931	-0.0409
σ	1.0530	0.7276	0.5038	0.8314	0.7026	1.1109	0.7843	0.8554	1.0803	0.5487	0.5291
a	1.1670	2.0313	3.8303	1.9440	2.2742	0.9718	1.8799	1.5514	1.2326	3.2875	3.4667
λ_+	1.0280	1.0384	1.0855	1.6044	1.0921	1.0000	1.0635	1.0540	0.9942	0.4083	0.9824
λ_-	1.0311	1.0786	1.1733	0.4052	1.0961	1.0000	1.0801	1.0925	1.6001	1.9144	1.2202
α	1.4717	1.6663	1.9189	1.2897	1.7461	1.3000	1.6610	1.5913	1.3256	1.5053	1.8437
A_2_MixedTS	0.0060	0.0055	0.0035	0.0065	0.0047	0.0055	0.0048	0.0046	0.0057	0.0061	0.0093
X_2_MixedTS	0.0400	0.0333	0.0345	0.0413	0.0365	0.0317	0.0363	0.0370	0.0393	0.0363	0.0474
A_1_MixedTS	0.0038	0.0036	0.0021	0.0042	0.0034	0.0037	0.0037	0.0038	0.0039	0.0034	0.0061
A_2_VG	0.0062	0.0066	0.0066	0.0071	0.0057	0.0092	0.0058	0.0048	0.0063	0.0075	0.0069
X_2_VG	0.0449	0.0346	0.0384	0.0435	0.0377	0.0345	0.0383	0.037	0.0385	0.0415	0.0555
A_1_VG	0.0042	0.0040	0.0041	0.00500	0.0042	0.0055	0.0044	0.0039	0.0042	0.0042	0.0047

Table 4.6: We fitted the MixedTS ang VG distribution to the empirical density of each sector and obtained the corresponding parameters for both models.

4.4 Risk Measures using the Mixed Tempered Stable distribution

4.4.1 Saddlepoint Approximation

In some cases we need to approximate the tail of the return density. Edgeworth expansions are frequently used to approximate distributions when the higher order moments are available. This methodology seems to work well in the center of distribution but not in the tails. In particular, it often produces negative values for densities in the tails. Saddlepoint expansion can be understood as an refinement of Edgeworth expansion on the tails. Wong [2008] proposes the saddle point technique to derive the small sample asymptotic distribution for the sample expected shortfall under a standard normal null hypothesis. In Cizek *et al.* [2011] is given a general introduction about how to approximate the density and the cumulative distribution function (cdf) of a continuous r.v. X whose moment generating function (mgf) $M_X(t)$ exists in an open set around zero. Here, we describe the main passages. We define the cumulant generating function (cgf)

$$K_X(t) = \ln[M_X(t)] \tag{4.24}$$

and the solution $\hat{s} = s(x)$ of the equation $x = K'_X(\hat{s})$ is the Saddle-Point at x for every $x \in$ *support*(X). The first order Saddle-Point approximation for the density $f_X(x)$ is:

$$\hat{f}_X(x) = \frac{1}{\sqrt{2\pi K''_X(\hat{s})}} \exp\{K_X(\hat{s}) - x\hat{s}\} \tag{4.25}$$

while the approximation for the cdf $F_X(x)$ obtained by Lugannani and Rice [1980] is :

$$\hat{F}_X(x) = \Phi(\hat{\omega}) + \phi(\hat{\omega}) \left\{ \frac{1}{\hat{\omega}} - \frac{1}{\hat{u}} \right\}, \quad x \neq E[X] \tag{4.26}$$

for $\hat{\omega} = \mathrm{sgn}(\hat{s})\sqrt{2\hat{s}x - 2K_X(\hat{s})}$, $\hat{u} = \hat{s}\sqrt{K''_X(\hat{s})}$. $\Phi(z)$ and $\phi(z)$ are the cdf and the density of a standard normal random variable. The condition $K'_X(0) = E[X]$ implies that $\hat{s} = 0$ is the saddlepoint for $E[X]$. Since $K_X(0) = 0$ and $\hat{\omega} = 0$ we can not use directly the expression for $x = E[X]$ but linear interpolation around the mean ensures continuity of the $\hat{F}_X(x)$. Second order approximation formula requires higher order derivatives of the cumulant generating function. Defining:

$$\hat{\kappa}_i = \frac{K^i_X(\hat{s})}{\left\{K''_X(\hat{s})\right\}^{\frac{i}{2}}} \tag{4.27}$$

and

$$a_i = \frac{\hat{\kappa}_4}{8} - \frac{5\hat{\kappa}_3^2}{24} \tag{4.28}$$

the second order saddle-point approximation formulas for the density and the cdf are respectively:

$$\begin{aligned} \tilde{f}(x) &= \hat{f}(x)(1+a_1) \\ \tilde{F}(x) &= \hat{F}(x) - \Phi(\hat{\omega})\left\{\hat{u}^{-1}\left(\frac{\hat{\kappa}_4}{8} - \frac{5}{24}\hat{\kappa}_3^2\right) - \hat{u}^{-3} - \frac{\hat{\kappa}_3}{2\hat{u}^2} + \hat{\omega}^{-3}\right\},\ x \neq E[X] \end{aligned} \tag{4.29}$$

The saddlepoint approximation can be used for computing some risk measures, in particular Martin [2006] gives an approximation for the integral appearing in the Expected Shortfall (ES) formula:

$$ES_\alpha(X) = \frac{1}{\alpha}\int_{-\infty}^{q} x f_X(x)dx \approx \frac{1}{\alpha}\left[\mu_X F_X(q) - f_X(q)\frac{q-\mu_X}{\hat{s}}\right] \tag{4.30}$$

Replacing the true density in (4.30) with the saddle-point first order approximation formula (4.25), the ES can be evaluated using the following first-order approximation formula:

$$\hat{ES}_\alpha(X) = \Phi(\hat{\omega}_q)\mu_X + \sqrt{\frac{1}{2\pi}}\exp\left\{-\frac{\hat{\omega}_q^2}{2}\right\}\left[\frac{\mu_X}{\hat{\omega}_q} - \frac{q}{\hat{u}_q}\right] \tag{4.31}$$

where $\hat{\omega}_q$, $\hat{u}_q$ are evaluated in $x = q$.
Using the second order saddle-point approximation (4.29) the corresponding Expected Shortfall approximated formula is:

$$\tilde{ES}_\alpha = \Phi(\hat{\omega}_q)\mu_X + \sqrt{\frac{1}{2\pi}}\exp\left\{-\frac{\hat{\omega}_q^2}{2}\right\}\left[\frac{\mu_X}{\hat{\omega}_q} - \frac{q}{\hat{u}_q} + \left(\frac{q}{\hat{u}_q^3} - \frac{\mu_z}{\hat{u}_q^3} - \frac{\mu_X}{\hat{u}_q^3} + \frac{q\hat{\kappa}_3}{2\hat{u}_q^2} - \frac{qa_1}{\hat{u}_q} - \frac{1}{\hat{s}_q\hat{u}_q}\right)\right] \tag{4.32}$$

These formulas can be easily handled in a parametric context and the obtained estimates are more stable than the historical quantities.

4.4.2 Numerical Analysis

In the previous sections we tested the ability of the new distribution to reproduce the features of the returns observed in the market. Now, we focus on some methodologies that allow the computation of risk measures in a parametric context in order to exploit the results obtained for the Mixed tempered Stable distribution. As a first exercise, we fit the marginal MixedTS directly to each GICS sector index empirical distribution and in table 4.7 we report the estimated parameters for each index.

	VFIAX	COND	CONS	ENRS	FINL	HLTH	INDU	INFT	MATR	TELS	UTIL
μ_0	-0.0011	0.0035	-0.0023	0.0023	0.0041	-0.0078	0.0073	0.0030	0.0046	0.0116	0.0064
μ	0.0021	-0.0016	0.0031	0.0000	-0.0019	0.0092	-0.0039	-0.0014	-0.0030	-0.0149	-0.0040
a	3.6321	9.0000	1.5784	9.0000	4.2702	7.2298	2.2836	1.6623	4.2977	9.0000	4.3842
σ	0.0045	0.0028	0.0062	0.0033	0.0047	0.0029	0.0069	0.0085	0.0048	0.0026	0.0040
λ_+	1.6337	1.9358	5.0716	1.8305	1.4665	6.8170	0.6007	1.8138	1.7252	2.1552	0.8187
λ_-	0.8290	2.0906	0.7012	2.0792	1.9103	0.7407	1.6886	4.4918	1.6860	1.9341	2.7837
α	0.7000	1.9900	0.7000	1.9900	1.4160	0.7000	0.7000	0.7000	1.6500	1.9900	0.7000

Table 4.7: Estimated Parameters for each Index in the dataset

We evaluate the cumulative distribution function using the formula that through the characteristic function $\phi_X(t)$ allows us to evaluate the cdf $F_X(x)$:

$$F_X(x) = \frac{1}{2} - \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{\left[e^{-itx}\phi_X(t)\right]}{it} \mathrm{d}t \tag{4.33}$$

We consider the VaR computed using the Inverse Fourier Transform and compare it with that obtained with historical simulation approach. The comparison for 5% and 10% levels of confidence is reported in Table 4.8

	$VaR_{0.05}^{MixedTS}$	$VaR_{0.05}^{Emp}$	$VaR_{0.1}^{MixedTS}$	$VaR_{0.1}^{Emp}$
VFIAX	0.0131	0.0130	0.0085	0.0092
COND	0.0118	0.0121	0.0086	0.0097
CONS	0.0118	0.0107	0.0072	0.0079
ENRS	0.0169	0.0166	0.0101	0.0112
FINL	0.0155	0.0154	0.0098	0.0105
HLTH	0.0128	0.0123	0.0082	0.0095
INDU	0.0118	0.0120	0.0108	0.0097
INFT	0.0169	0.0145	0.0119	0.0111
MATR	0.0154	0.0168	0.0110	0.0140
TELS	0.0142	0.0129	0.0104	0.0104
UTIL	0.0118	0.0115	0.0085	0.0089

Table 4.8: In this table we compare the empirical VaR with the theoretical one obtained for the ten GICS and the VFIAX fund index.

For each index, we compare the empirical ES estimates with the estimates using four different methodologies.

1. Numerical integration of the expected value, i.e the distribution is obtained through the Inverse Fourier Transform. We recall that, under the assumption of the existence for the $E(X)$ [1], the Expected Shortfall can be written as (see Cizek *et al.* [2011]):

$$ES_\alpha(X) := E[X|\ X \leq x_\alpha] = x_\alpha - \frac{1}{\alpha}\int_{-\infty}^{x_\alpha} F(u)\,\mathrm{d}u \tag{4.34}$$

2. The second approach is based on Monte Carlo simulation. The random number generator is built using the Inverse Transform Sampling method (see Glasserman [2003]) based on the following two steps:

   a Generate a random number $\hat{u}$ from a uniform distribution $U([0,1])$.

   b Find x such that the equation $F_X(x) = \hat{u}$.

3. First order Saddle Point Approximation formula.

4. Second order Saddle Point Approximation formula.

Tables 4.9 and 4.10 report the ES computed with the four methodologies and in the last column we have the empirical ES obtained using the historical simulation. In this case the windows size is the last year of observations. In table 4.9, we have computed the ES at 5% level of confidence, while in table 4.10 the ES is at 10% level of confidence.

[1] For the Mixed Tempered Stable distribution, this condition is ensured by the existence of moment generating function. The condition of existence of the first moment is less obvious when we obtain the Geo Stable as a special case.

	$ES_{0.05}^{MixedTS}$	$ES_{0.05}^{Fourier}$	$ES_{0.05}^{SPA1}$	$ES_{0.05}^{SPA2}$	$ES_{0.05}^{Emp}$
VFIAX	0.0173	0.0179	0.0205	0.0204	0.0170
COND	0.0157	0.0158	0.0155	0.0164	0.0170
CONS	0.0176	0.0177	0.0200	0.0196	0.0163
ENRS	0.0184	0.0185	0.0182	0.0192	0.0230
FINL	0.0192	0.0194	0.0194	0.0202	0.0206
HLTH	0.0177	0.0188	0.0209	0.0204	0.0171
INDU	0.0210	0.0210	0.0235	0.0233	0.0197
INFT	0.0209	0.0206	0.0231	0.0223	0.0191
MATR	0.0215	0.0211	0.0212	0.0218	0.0233
TELS	0.0186	0.0188	0.0189	0.0192	0.0216
UTIL	0.0161	0.0161	0.0165	0.0165	0.0170

Table 4.9: In this table we give the results of computing the ES for $\alpha = 0.05$

	$ES_{0.1}^{MixedTS}$	$ES_{0.1}^{Fourier}$	$ES_{0.1}^{SPA_1}$	$ES_{0.1}^{SPA_2}$	$ES_{0.1}^{Emp}$
VFIAX	0.0142	0.0147	0.0161	0.0166	0.0140
COND	0.0128	0.0129	0.0123	0.0141	0.0139
CONS	0.0132	0.0139	0.0158	0.0159	0.0126
ENRS	0.0153	0.0152	0.0145	0.0165	0.0184
FINL	0.0154	0.0155	0.0151	0.0173	0.0166
HLTH	0.0133	0.0145	0.0154	0.0156	0.0139
INDU	0.0165	0.0175	0.0183	0.0195	0.0153
INFT	0.0168	0.0183	0.0199	0.0198	0.0160
MATR	0.0172	0.0170	0.0168	0.0184	0.0191
TELS	0.0152	0.0155	0.0153	0.0160	0.0163
UTIL	0.0129	0.0131	0.0132	0.0140	0.0133

Table 4.10: In this table we give the results of computing the ES for $\alpha = 0.10$

The last comparison based on marginal distribution regards the expectile based risk measure. As observed in Chapter 2 a plausible value of the weight parameter for the considered dataset is 0.012. The parametric expectile is obtained by using the second approach described for the ES computation, i.e by using Monte Carlo technique.
The results are reported in Table 4.11

	$e_{0.012}^{MixedTS}$	$e_{0.012}^{Emp}$
VFIAX	0.0135	0.0136
COND	0.0127	0.0134
CONS	0.0130	0.0132
ENRS	0.0141	0.0184
FINL	0.0156	0.0162
HLTH	0.0152	0.0135
INDU	0.0156	0.0157
INFT	0.0183	0.0152
MATR	0.0173	0.0188
TELS	0.0156	0.0175
UTIL	0.0137	0.0139

Table 4.11: We fix $\tau = 0.012$ coherently with the value obtained in chapter 2 for the VFIAX fund and then compare the parametric and non parametric results.

Chapter 5

Conclusions and Future Research

In this thesis we investigate some methodologies applied to risk attribution analysis. Based on the existing literature, we emphasize the problems encountered in their implementation and proposed some modifications that were supported by the empirical analysis.

In the first three chapters we focus on non parametric approaches and study several methods for identifying the factors that most influence asset or portfolio returns. In Chapter 2 from the empirical analyisis performed, it seems that τ can be defined as an increasing function of the kurtosis but further investigation is needed in order to model the weight parameter. In Chapter 3, starting from the seminal work of Menchero and Poduri [2008], we obtain a model through which we are able to decompose the return vector in three uncorrelated quantities: the user defined factors, the projected factors and the idiosyncratic error term. It is easy to implement and each risk component has a financial interpretation. In the future, we will apply this methodology to more complex portfolios and for a higher number of custom risk factors considered.

In the last chapter, we propose a new distribution, namely Mixed Tempered Stable. The flexibility of this distribution, supported from our empirical analysis, derives from its capability of having as special cases distributions that are able to fit the tails.

The study of risk attribution using the MixedTS will be object of future research. The definitions of modified VaR Zangari [1996] and modified ES Boudt *et al.* [2009] that use asymptotic expansions can be a starting point. Another interesting argument is the study of efficient algorithms that generate sparse linear models as the one recently introduced in Zhang *et al.* [2009]. In a parametric context, with a large number of factors considered, they facilitate numerical implementation of parameters estimation.

Bibliography

P. Abad and S. Benito. Value at risk in fixed income portfolios: a comparison between empirical models of the term structure. *Business Review Cambridge*, 7(2):342, 2007.

C. Acerbi. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518, July 2002.

P. Artzner, F. Delbaen, J. Eber, and D. Heath. Coherent measures of risk. *Mathematical Finance*, 9(3):203–228, 1999.

W. Aussenegg and T. Miazhynskaia. Uncertainty in value-at-risk estimates under parametric and non-parametric modeling. *Financial Markets and Portfolio Management*, 20(3):243–264, 2006.

A.D Back and A. S. Weigend. A first application of independent component annalysis to extracting structure from stocks returns. *International Journal of Neural System*, 8:473–484, 1997.

F. Bellini and V. Bignozzi. Elicitable risk measure. *Working paper*, 2013.

F. Bellini and E. Salinelli. Independent component analysis and immunization: An exploratory study. *International Journal of Theoretical and Applied Finance*, 6(7):721–738, 2003.

F. Bellini, B. Klar, A. M. Muller, and E. Rosazza Gianin. Generalized quantiles as risk measures. *Working paper*, 2013.

BIS. Fundamental review of trading book capital requirements: consultation by the basel committee. 2012.

F. Black. Capital market equilibrium with restricted borrowing. *Journal of Business*, 45:444–454, 1972.

T. Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):397–327, 1986.

K. Boudt, B. H. Peterson, and C. Croux. Estimation and decomposition of downside risk for portfolios with non-normal returns. *Journal of Risk*, 11(2):79–103, 2009.

G. P. Brinson. Measuring non-us equity portfolio performance. *Journal of Portfolio Management*, pages 73–76, 1985.

Mark M. Carhart. On persistence in mutual fund performance. *Journal of Finance*, 52(1):57–82, 1997.

P. Carr and L. Wu. Time-changed levy processes and option pricing. *Journal of Financial Economics*, 71(1):113–141, 2004.

P. Carr, H. Geman, D.B. Madan, and M. Yor. The fine structure of asset returns: an empirical investigation. *Journal of Business*, 75(2):305–332, 2002.

Y. Chen, W. Hardle, and V. Spokoiny. Ghica risk analysis with gh distributions and independent components. *Journal of Empirical Finance*, 17:255–269, 2010.

P. Christoffersen. Encyclopedia of quantitative finance-backtesting. *John Wiley and Sons, Second Edition*, 2010.

P. J. Cizek, W. Hardle, and R. Weron. Statistical tools for finance and insurance. *Second Edition, Springer*, 2011.

P. Comon. Independent component analysis, a new concept? *Signal Process.*, 36(3):287–314, 1994.

R. Cont and P. Tankov. Financial Modelling with Jump Processes. *Chapman & Hall/CRC Financial Mathematics Series*, II, 2003.

R. Cont. Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.

B. Denis, C. Jason, L. Feifei, and S. Omid. Risk parity portfolio vs. other asset allocation heuristic portfolios. *Journal of Investing*, 20(1):108–118, 2011.

E. Falkenstein and J. J. Hanweck. Minimizing basis risk from non parallel shifts in the yield curve part ii: principal components. *Journal of Fixed Income*, 7(1):85, 1997.

E.F. Fama and K. French. The cross-section of expected stock returns. *Journal of Finance*, pages 427–466, June 1992.

E.F. Fama and K. French. Size and book to market factors in earnings and returns. *Journal of Finance*, (50):131–155, 1995.

Paul Glasserman. Monte Carlo Methods in Financial Engineering (Stochastic Modelling and Applied Probability) (v. 53). *Springer*, 2003.

T. Gneiting. Making and evaluating point forecasts. *Journal of the American Statistical Association*, (106):746–762, 2011.

G. H. Golub and C. Van Loan. Matrix computations (3rd ed.). *Baltimore: Johns Hopkins University Press*, 1996.

C. Gourieoux, J. Laurent, and Olivier S. Sensitivity analysis of value at risk. *Journal of Empirical Finance*, 7:225–245, 2000.

A. Hyvarinen, J. Karhunen, and E. Oja. Independent component analysis. 2001.

A. Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 10:626–634, 1999.

I. T. Jolliffe. Principal component analysis. 2002.

M.C. Jones. Expectiles and m-quantiles are quantiles. *Statistics and Probability Letters*, 20(2):149–153, 1994.

R. M. Kalman. A new approach to linear filtering and prediction problems. *Journal of Basic Enginering-Tansaactions of the ASME. Series D*, 82:35–45, 1960.

Y.S. Kim, S. T. Rachev, M. L. Bianchi, and F.J. Fabozzi. Financial market models with lÃľvy processes and time-varying volatility. *Journal of Banking and Finance*, 32(7):1363–1378, 2008.

I. Koponen. Analytic approach to the problem of convergence of truncated lévy flights towards the gaussian stochastic process. *Physical Review E*, 52:1197âĂŞ1199, 1995.

T. J. Kozubowski, K. Podgórski, and G. Samorodnitsky. Tails of lévy measure of geometric stable random variables. *Cornell University Operations Research and Industrial Engineering*, 1997.

T. Kozubowski, K. Podgorski, and G. Samorodnitsky. Tails of levy measure of geometric stable random variables. *Working paper*, pages 1–3, 2011.

C. Kuan, Yeh J., and Y. Hsu. Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics*, 150(2):261–270, 2009.

C. M. Kuan, J. H. Yeh, and Y. C. Hsu. Assessing value at risk with care, the conditional autoregressive expectile models. *Journal of Econometrics*, 150(2):261–270, 2009.

U. Küchler and S. Tappe. Bilateral gamma distributions and processes in financial mathematics. *Stochastic Processes and their Applications*, 118(2):261–283, 2008.

U. Küchler and S. Tappe. On the shapes of bilateral gamma densities. *Statistics and Probability Letters*, 78(15):2478–2484, 2008.

U. Küchler and S. Tappe. Option pricing in bilateral gamma stock models. *Statistics and Decisions*, 27:281–307, 2009.

U. Küchler and S. Tappe. Tempered stable distributions and processes. *Stochastic Processes and their Applications*, 2013. in press.

J. Lintner. The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47:13–37, 1965.

F. A. Longstaff, Pan J., L.H. Pedersen, and K.J. Singleton. How sovereign is sovereign credit risk? *American Economic Journal: Macroeconomics*, 3(2):75–103, 2011.

A. Loregian, L. Mercuri, and E. Rroji. Approximation of the variance gamma model with a finite mixture of normals. *Statistic & Probability Letters*, 82(2):217–224, 2012.

C. Lu, T. Lee, and C. Chiu. Financial time series forecasting using independent component analysis and support vector regression. *Decision Support System*, 47:115–125, 2009.

R. Lugannani and S. Rice. Saddlepoint approximation for the distribution of the sum of independent random variables. *Adv. in Appl. Probab.*, 12:475–490, 1980.

D.B. Madan and B. Seneta. The vg model for share market returns. *Journal of Business*, 63:511–524, 1990.

D.B. Madan and E. Seneta. The variance gamma (v.g.) model for share market returns. *Journal of Business*, 63:511–524, 1990.

D. Madan. Equilibrium asset pricing with non-gaussian returns and exponential utility. *Quantitative Finance*, 6(6):455–463, 2006.

D. B. Madan. Equilibrium asset pricing: with non-gaussian factors and exponential utilities. *Quantitative Finance*, 6(6):455 – 463, 2006.

B. B. Mandelbrot. The variation of certain speculative prices. *The Journal of Business*, 36, 1963.

M. Marchioro and D. Borrello. Fast risk decomposition. *Quantitative Research Series. StatPro website*, 2013.

H. Markowitz. Portfolio selection. *Journal of Econometrics*, 7(1):77–91, 1952.

R. Martin. The saddlepoint method and portfolio optionalities. *Risk Magazine*, 19:93–95, 2006.

J. Menchero and V. Poduri. Custom factor attribution. *Financial Analysts Journal*, 64(2):81–92, 2008.

A. Meucci. Risk contributions from generic user-defined factors. *Risk*, 20(6):84–88, 2007.

A. Meucci. Review of linear factor models. *http://ssrn.com/abstract=1635495*, 2010.

J. Moody and L. Wu. High frequency foreign exchange rates: price behaviour analysis and true price models. *Nonlinear Modeling of High Frequency Financial Times Series*, pages 23–47, 1998.

W. K. Newey and J. L. Powell. Asymmetric least squares estimation and testing. *Econometrica*, 55:819–847, 1987.

K. H. Osband. Providing incentives for better cost forecasting. *Ph.D. thesis , University of California, Berkeley.*, 1985.

K. Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine*, 2:559–572, 1901.

H. Reisman and G. Zohar. Short-term predictability of the term structure. *Journal of Fixed Income*, 14(3):7–15, 2004.

RiskMetrics. Technical document fourth edition. *J. P. Morgan Reuters*, December 1996.

R. T. Rockafellar and S. Uryasev. Optimization of conditional value-at-risk. *Journal of Risk*, 2:21–41, 2000.

K. I. Sato. Lévy Processes and Infinitely Divisible Distributions. *Cambridge studies in advanced mathematics*, II, 1998.

L.J. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):419–441, 1971.

0. Scaillet. Nonparametric estimation and sensitivity analysis of expected shortfall. *Mathematical Finance*, 14(1):74–86, 2002.

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

W.F. Sharpe. Budgeting and monitoring pension fund risk. *Financial Analysts Journal*, 58(5):74–86, 2002.

P. Shepard. The barra integrated model (bim 301). *Research notes MSCI*, 2011.

G. M. Soto. Duration models and irr management: a question of dimensions? *Journal of Banking and Finance*, 28(5):1089–1110, 2004.

D. Straumann and T. Garidi. Developing an equity factor model for risk. *RiskMetricsGroup*, pages 89–128, 2011.

D. Tasche. Risk contributions and performance measurement. *Working paper, Technische Universitat München*, 1999.

D. Tasche. Expected shortfall and beyond. *Journal Banking and Finance*, 26(7):1519–1533, 2002.

M. Tweedie. An index whichg distinguishes between some important exponantial families. In *Proc. Indian Statistical Institute Golden Jubilee International Conference,*, pages 579–604. J. Ghosh and J. Roy (Eds.), 1984.

K. Wong. Backtesting trading risk of commercial banks using expected shortfall. *Journal of Banking and Finance*, 32(7):1404–1415, 2008.

E. Wu, P. Yu, and W. Li. Value at risk estimation using indep endent com- p onent analysis-generalized autoregressive conditional heteroscedasticity (ica- garch) models. *International Journal of Neural Systems*, 16:371–389, 2006.

Y. Yamai and T. Yoshiba. Value-at-risk versus expected shortfall:a practical perspective. *Journal Banking and Finance*, 29:997–1015, 2005.

P. Zangari. A var methodology for portfolios that include options. *RiskMetrics Monitor*, pages 4–12, 1996.

K. Zhang, H. Peng, L. Chan, and A. Hyvarinen. Ica with sparse connections: Revisited. *Lecture Notes in Computer Science*, 5441:195–202, 2009.