Università
Ca'Foscari
Venezia

Corso di Laurea Magistrale
in Economia e Finanza

Tesi di Laurea

Ca' Foscari
Dorsoduro 3246
30123 Venezia

Bayesian Stochastic Volatility Analysis for Hedge Funds

Relatore
Ch. Prof. Roberto Casarin

Laureando
Alsu Salikhova
Matricola 845705

Anno Accademico
2012 / 2013

ABSTRACT

Many daily and weekly financial series exhibit time-varying volatility and volatility clustering. Stochastic volatility models are a way to model these stylized facts. The presence of the latent components in the volatility model carries difficulties in the application of maximum likelihood methods and call for the use of Bayesian inference. In this thesis, we introduce a new Bayesian multivariate stochastic volatility model and propose a suitable inference procedure. We discuss the choice of priors and apply a MCMC simulation algorithm for estimating the parameters and the latent variables. We provide an application of the model and of the MCMC algorithm to hedge funds data.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to Professor Roberto Casarin, my thesis supervisor, for his patient guidance, enthusiastic encouragement and useful critiques of this thesis. I just could not wish for a better or friendlier supervisor.

I wish also to thank Professor Mario Bellia and Michele Trova from Veneto Banca bank for their advice in hedge fund data collection.

I would like to extend my thanks to Carlo Miclet for his technical assistance in downloading the data.

A special thanks to Professor Monica Billio, my exchange program coordinator, for her priceless advice, attention, support and encouragements during all my exchange program period in Ca'Foscari University of Venice. Without her this thesis would never existed.

Finally, I wish to thank my beloved mother for her love, support and faith in me and everything I do.

TABLE OF CONTENTS

Abstract ii
Acknowledgements iii
Table of contents iv
List of tables v
List of figures v
Introduction 1
Chapter 1 – Hedge Funds 3
1.1. History, Definition and Role of Hedge Funds 3
1.2. Data and Strategy Classification 5
1.3. Stylized Facts 8
Chapter 2 – Stochastic Volatility 16
2.1. Introduction 16
2.2. Bayesian inference 16
2.3. Markov Chain Monte Carlo Simulation Method 17
2.4. Gibbs Sampling 18
2.5. Metropolis Algorithm 19
2.6. Bayesian Univariate Stochastic Volatility Model 20
2.5. Hierarchical Multivariate Stochastic Volatility Model 25
Chapter 3 - Simulation and Application to the Hedge Funds 32
3.1. Simulation Results 32
3.2. Empirical Application to Hedge Fund Returns 34
Chapter 4 - Conclusion 40
Appendix 41
References 42

LIST OF TABLES

Table 1.1: Summary statistics for individual hedge fund returns 9
Table 1.2: Summary statistics of groups of hedge fund returns.....13

LIST OF FIGURES

Figure 1.2: Number of different strategy groups. 4
Figure 1.3: Mean and standard deviation of returns of 165 individual hedge funds and three benchmarks.....10
Figure 1.4: Kurtosis and skewness of returns of 165 individual hedge funds and three benchmarks.....10
Figure 1.5: Volatility vs. kurtosis and skewness of returns of 165 individual hedge funds and three benchmarks.....11
Figure 1.6: Kurtosis vs. skewness of 15 classes of hedge funds and three market proxies....13
Figure 1.7: Volatility vs. asymmetry of 15 classes of hedge funds and three market proxies14
Figure 1.8: Volatility vs. kurtosis of 15 classes of hedge funds and three market proxies 14

Figure 2.1: Representation of Hierarchical Structure and setting of the prior parameters 27

Figure 3.1: Simulated trajectories of 250 observations from the observable process $y_{t,j}$ 32
Figure 3.2: MCMC raw output for the first latent variables $h_{t,104}$ 33
Figure 3.3: True parameters α_j , φ_j and the estimated parameters $\hat{\alpha}_j$, $\hat{\varphi}_j$.....34
Figure 3.4: Up: Log-return of 165 hedge funds. Bottom: Square of log-return of 165 hedge funds 34
Figure 3.5: MCMC raw output for the latent variables $\hat{h}_{t,85}$ 35
Figure 3.7: True and estimated values of intercept $\alpha_{t,j}$ and persistence of volatility $\varphi_{t,j}$.....37
Figure 3.8: Estimated values of intercept $\hat{\alpha}_{t,j}$ and persistence of volatility $\hat{\varphi}_{t,j}$according to different management styles.....37
Figure 3.9 a): Standard deviation and kurtosis of individual hedge funds 38
Figure 3.9 b): Estimated values of intercept $\hat{\alpha}_{t,j}$ and persistence of volatility $\hat{\varphi}_{t,j}$.....38
Figure 3.10: Zoom on the Figure 9b).....39

Introduction

The analysis of dynamic phenomena is a common problem to many fields such as physics, statistics, engineering, economics and finance. Modelling changes in financial variables represents an important issue in econometrics. The use of stochastic models such as dynamic models brings the application of economics and finance closer to the empirical reality, and allows to describe more accurately many peculiarities of the financial variables and to make better decisions in a situation of uncertainty.

Many daily and weekly financial time series exhibit time variations in the conditional velocity. One of the most popular ways to model time-varying velocity is the auto regressive conditional heteroskedasticity (ARCH) model introduced by Engle (1982), or one of its variants (e.g., GARCH, Bollerslev (1986), EGARCH, Nelson (1991)). ARCH models are attractive because they offer a parsimonious model for the dynamics of the volatility, and are easy to estimate. A major drawback of ARCH models is that they do not always manage fully reproduce the characteristics of financial series (Bauwens, Lubrano, and Richard, (1999)). ARCH models have some limitations: with a single error term they have to deal with the error of observation for the mean and with the dynamics of the volatility.

A stochastic volatility (SV) model is an alternative approach where the variance is modelled as an *unobserved (or latent)* component. Interest in SV models dates back to 1973, when Clark (1973) proposed a mixture model for the distribution of stock returns. His seminal paper showed that it is possible to move away from Brownian motion assumption, also avoiding the use of heavy-tail processes viewed as stable processes, which could be empirically unattractive. Since 1973 many studies of dynamic models have been done and a variety of procedures have been proposed for estimating the SV model: method of moments (Duffie and Singleton (1993); Andersen and Sorensen (1996)), efficient method of moments (Bansal et al. (1993); Gallant et al. (1997)), simulated likelihood (Danielson (1994)).

Here, we employ the Bayesian methods for inference and prediction in a simple class of SV models proposed by Jaquier et al. (1994), in which the logarithm of the conditional volatility follows an autoregressive dynamics.

The presence of latent factors in time series analysis increases the degree of complexity, namely it could be hard to derive the likelihood for parameters estimation. Hence, the Bayesian Inference approaches may be a solution of a statement about an

uncertain future. Bayesian Inference is a method of inference in which Bayes' rule is used to update the probability estimate for a hypothesis as additional evidence is learned. This approach has several advantages. In the Bayesian Inference approach the prior knowledge of a previous model can be used to inform the current model. Knowledge is available in several forms: historical information and professional wisdom or expertise. The uncertainty is included in the probability model, so predictions about the parameters are more realistic. The use of MCMC simulation algorithm during estimation allow to handle complicated models, for instance hierarchical models, the MCMC algorithm is unbiased with respect to sample size and can provide realistic estimations even for small samples.

The aim of this thesis is to introduce a new Bayesian hierarchical stochastic volatility model, and to propose a solution to the related inference problems. We will show how simulation methods such as Gibbs sampling and Metropolis-Hastings algorithms apply to our stochastic volatility model.

Another goal of the thesis is the investigation of clustering phenomena in volatility for a sample of hedge fund returns. We will verify whether volatility clustering is similar, in terms of fund specific parameters, across the hedge funds.

The rest of the paper is outlined as follows. Chapter 1 introduces the general description of the hedge funds and their performance. Chapter 2 provides with representation of SV model in a Bayesian framework. The chapter reviews MCMC algorithm such as the Gibbs sampler algorithm and Metropolis-Hastings algorithm and introduces a hierarchical multivariate SV model. Chapter 3 applies the model to analysis of return data and provides robustness check. Chapter 4 concludes.

Chapter 1 – Hedge Funds

1.1. History, Definition and Role of Hedge Funds

According to the latest HFR Global Hedge Fund Industry Report, released on April, 19th 2013 by HFR the *hedge fund* sector is one of the fastest growing sectors of the financial services industry, which is currently estimated at $2.375 trillion (HFR Global Hedge Fund Industry Report - First Quarter 2013). In the twenty years, the hedge fund industry has grown from 300 funds in 1990 to more than 10000 funds today (HedgeCo.Net). One of the main reasons for such interest is the performance characteristics of hedge funds: many hedge funds can provide double-digit returns for their investors. This happens because many hedge funds managers take simultaneously long and short positions, thus investors profit from both positive and negative information. Figure 1 shows return performance of hedge funds compared to stocks, bonds and commodities.

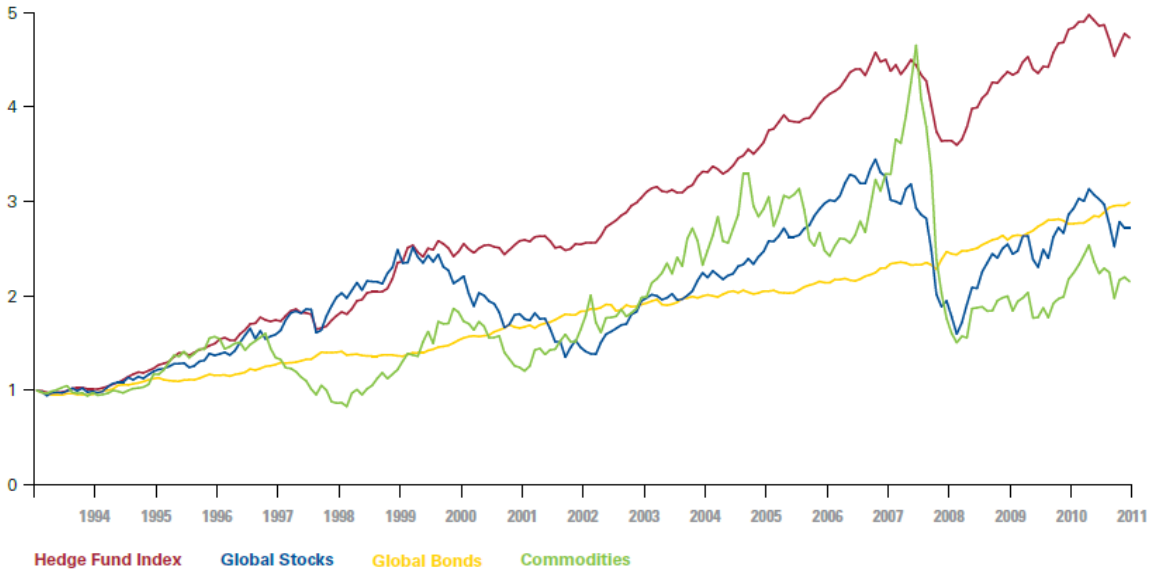


Figure 1.1: Hedge fund returns compared to stocks, bonds and commodities
Source: KPMG International, 2012.

Historically, the first hedge fund was created in 1949 by Alfred W. Jones. By using leverage and short selling, he effectively "hedged" risk in the marketplace. His hedge fund greatly outperformed mutual funds of that time. However, hedge funds really got interest in the 60's. Warren Buffet and George Soros were the first to take an interest in Jones' strategy, and over the next years, 130 hedge funds were born.

Hedge funds, like other alternative investments such as real estate and private equity, are thought to provide returns that are uncorrelated with traditional investments. This attracted an increasing number of individual and institutional investors. Hedge fund investment strategies vary from one to another. Every hedge fund is asked to classify itself into one of a number of different strategy groups depending on the main type of strategy followed. There is about 20 different hedge fund strategies used by hedge fund managers today. Figure 2 shows the different strategies of 4526 hedge funds depending on managing strategy.

We remark that many hedge funds are not "hedged" at all. In the sake of large returns hedge funds managers often take highly speculative positions. Hedge funds were never required to report to regulators any changes in trading and strategies, unlike mutual funds. Due mainly to this fact, it is difficult to assess and define exactly the asset allocation, concentrations of different strategies and the changes in historical prices.

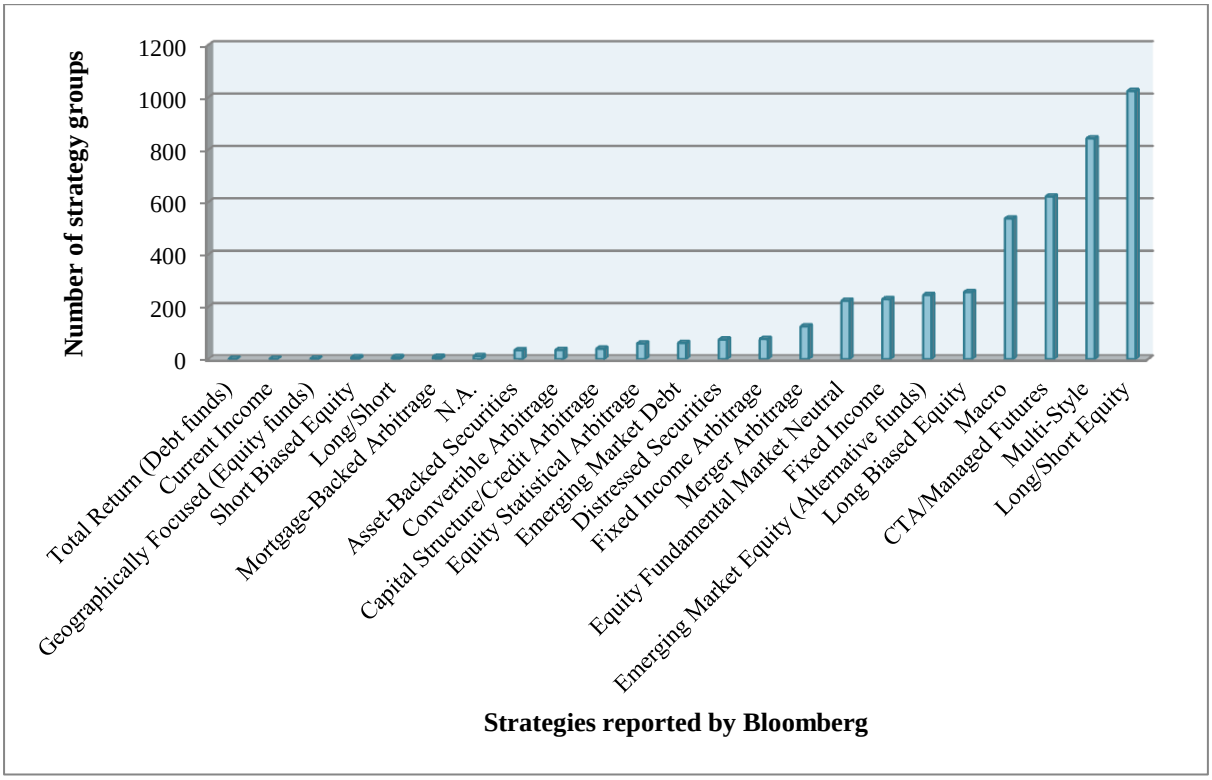


Figure 1.2: Number of different strategy groups depending on the main type of strategy followed. Frequency: bi-weekly, daily, monthly, weekly and N.A.
Source: Data from *Bloomberg.*

1.2. Data and Strategy Classification

The tremendous growth and popularity of hedge funds has attracted attention of a lot of researchers who have been concentrated to study all aspect of these relatively new investment vehicles. Fung and Hsieh (2001) and Mitchell and Pulvino (2001) analyzed the risks in hedge fund styles, Fung and Hsieh (1997) contributed to the performance and risk exposure analysis, Getmansky et al. (2004) provided an econometric analysis of liquidity in hedge funds returns. Recent contributions to the analysis of risk exposure of hedge funds belong to Billio et al. (2009) and Billio et al. (2012), Klaus and Rzepkowski (2009).

In this chapter we will examine statistical properties of hedge fund returns and we will represent some performance features of returns.

The data used in this study were obtained from Bloomberg Professional® platform, which is the one of the best known and largest databases currently available for professional investors. As of May, 16th 2013, the database at our disposal contained weekly net asset value on a total of 676 hedge funds. The currency of the Net Asset Value of weekly hedge funds is available in US dollars. Hedge funds are privately offered, pooled investment vehicles and are not widely available to the public and many hedge fund series among 676 weekly available contained missing data. In order to replace missing data we filled missing values with the mean of valid surrounding values. The span of nearby points is the number of valid values above and below the missing value used to compute the mean.

For the empirical analysis in this thesis, we use net asset value of 165 Hedge Funds from Bloomberg database from 04/01/2008 to 14/12/2012, overall of 259 observations. We restricted our attention to 165 hedge funds to have a larger sample of observations for each hedge fund. We shall note that our database contains information only on existing funds and it is affected by potential bias – a survivorship bias (Fung and Hsieh (2000)). Survivorship bias occurs when hedge funds are excluded from a database because they have gone out of business for any number of reasons, including bankruptcies, liquidations, mergers, name changes, and voluntary stoppage of information reporting or just no longer report due to the poor results. Only "surviving" funds left, funds that are still in operation and are reporting information to the database vendor at the end of the data sample.

The selected 165 hedge funds are grouped into 15 different strategy classes, among them 11 classes are equity-based-strategies and 4 classes which are non-equity-based-strategies (See Fung and Hsieh (2000), Billio et al.(2009), Trang (2006)):

Equity-based-strategies:

1. ***Capital Structure/Credit Arbitrage***: Funds that are going long one security in a company's capital structure while at the same time going short another security in that same company's capital structure.
2. ***Distressed/High Yield Securities:*** Funds invest in securities of companies of some financial distress such as reorganizations, bankruptcies, distressed sales, and other corporate restructurings. The securities range from senior secured debt to common stock.
3. ***Emerging Market Debt*****:** Funds that invest in bonds issued by less developed countries.
4. ***Emerging Market Equity:*** Funds that specialize in investment in the securities of emerging market countries.
5. ***Equity Market Neutral:*** Funds that seek to generate returns from taking simultaneously long and short position of the same size within the same market while neutralizing the systematic risk of the market.
6. ***Equity Statistical Arbitrage:*** Funds that seek to generate returns from statistical mis-pricing of one or more assets based on the expected value of these assets.
7. ***Long-Biased Equity:*** Funds that maintain a net long exposure to the market (compared to short positions).
8. ***Long/Short Equity*****:** Funds that invest on both the long and the short side of the equity market. These funds seek to reduce the overall market exposures of the hedge fund portfolio. Returns can be generated from both long and short sides.
9. ***Mortgage-Backed Arbitrage:*** Funds that seek pricing inefficiencies in the primarily US-based mortgage-backed securities market.
10. ***Multi-Style:*** Funds that seek to profit from using a variety of different strategies and adjusting their allocations based upon perceived opportunities.
11. ***Short-biased****:* Funds that seek to profit primarily from the short sale of securities, but may buy securities as a hedge.

Non-equity-based-strategies:

1. ***CTA/Managed Futures*** (CTA: Commodity Trading Advisor): Funds that go long or short in futures contracts in areas such as metals (gold, silver), grains (soybeans, corn, wheat), equity indexes (S&P futures, Dow futures, NASDAQ 100 futures), soft commodities (cotton, cocoa, coffee, sugar) as well as foreign currency and U.S government bond futures.

2. ***Fixed Income:*** Funds using strategies in which the investment thesis is predicated on realization of a spread between related instruments in which one or multiple components of the spread is a fixed income.
3. ***Fixed Income Arbitrage:*** Funds that seek to profit from mispricing among similar fixed income securities.
4. ***Global Macro:*** Funds go long and short in stocks, bonds, currencies, and derivatives, including options and futures. These funds use leverage. The best known fund is Quantum Fund of George Soros.

One of the distinguishing features of hedge funds from traditional long-only stock and bond investing is a search for positive returns in up, down and tranquil market conditions. The market in this context is the equity (stock) and bond market. That is why it is reasonable to use as market proxy indices representing the performance of stock market and bond market with weekly closing quotes. As a stock market proxy we use S&P 500 index. In the sake of the larger sample of observations as possible we did not eliminate non-equity-based strategies, such as Fixed Income, Fixed Income Arbitrage, Macro and CTA/Managed Futures. Fixed income securities do not trade on open exchange, and bond prices are therefore less transparent. Benchmarks for non-equity-based strategies are created by large broker-dealers services providers that buy and sell bonds, for instance J.P. Morgan. That is why including in our analysis JPM GABI as bond market proxy in relevant. We also assume that hedge funds are investing all over the world, in Europe for instance, and there is an evidence of including in our comparative analysis the third index, which represents performance in European stock market, that is Morgan Stanly Capital International Europe (MSCI) index. These three indices will serve us as benchmarks to analyze how well the hedge funds are performing.

S&P 500 index is a capitalization-weighted index of 500 stocks. The index is designed to measure performance of the broad US economy through changes in the aggregate market value of 500 stocks representing all major industries. The MSCI (Morgan Stanly Capital International) Europe index is a free-float weighted equity index designed to measure the equity market performance of the developed markets in Europe. JPM GABI is a fixed income global benchmark. It consists of the JPM GABI US, a U.S. dollar denominated, investment-grade index spanning asset classes from developed to emerging markets, and the JPM GABI extends the U.S. index to also include multi-currency, investment-grade instruments.

1.3. Stylized Facts

In this subsection we analyze statistical properties of hedge fund returns. We examine risk and return relationship and statistical characteristics of return distributions. We focus on whether different individual hedge funds and different classes of hedge fund show dissimilarities in terms of normality or deviation from normality.

In this part we find that the deviation from normality of hedge funds returns exists at the weekly level for all (with some minor exceptions) different indices and strategies. The deviation from normality is measured through skewness, kurtosis and Jarque-Bera test.

Table 1.1, Figure 1.3 and Figure 1.4 describe the sample size, mean, standard deviation, skewness and kurtosis for weekly returns of 165 individual hedge funds as well as for the S&P 500, MSCI Europe and JPM GABI indices. The distribution of individual hedge funds returns is leptokurtic and exhibits a significant degree of skewness relative to the Normal distribution. From the Table 1.1 we learn that the distribution of all hedge funds and all three indices exhibit an excess of kurtosis. That means that there is a higher than normal probability of big positive and negative return realizations and this calls for the case of stochastic volatility models. The distributions of the big majority of individual hedge funds return and two indices S&P500 and MSCI Europe exhibit negative skewness. A distribution with negative skew has greater probability of below-average returns than the probability for above-average returns. The more the skew is negative, the greater is the probability of below-average returns. We implement a Jarque-Bera test, which represents a goodness-of-fit measure of departure from normality, based on the sample kurtosis and skewness. The *p*-value of Jarque-Bera test for almost all hedge funds and for three indices is below the significance level of 5% or equal to zero, and the test rejects the null hypothesis that the distribution is normal. For one hedge fund in Multi-Style Management Strategy and one hedge fund from Emerging Market Debt class the *p*-value of Jarque-Bera test is greater than 5%. That means that the distribution of those hedge funds is Normal. We notice the positive skew of the distribution for the individual hedge funds from different management style classes: CTA/Manages Futures (27% of the class sample), Emerging Market Debt (100% of the class sample), Emerging Market Equity (18% of the class sample), Equity Fundamental Market Neutral (4,5% of the class sample), Equity Statistical Arbitrage (66,7% of the class sample), Long/Short Equity (6% of the class sample), Macro (25% of the class sample), Multi-Style (6% of the class sample), and JPM GABI as well.

Our investigation of outliers shows that, only two individual hedge funds from CTA/Managed Futures and Fixed Income Arbitrage outperform all three market proxies. They provide a relative high return and at the same time they are relatively less risky. The most risky hedge funds are one from Emerging Market Equity class of hedge funds and one form Equity Fundamental Market Neutral. They underperform three benchmarks.

Table 1.1: Summary statistics for individual hedge fund returns

Strategy	N	Excess of kurtosis	Negative skewness	JB *p*-value
Capital Structure/Credit Arbitrage	5 (3%)	√	√	0.00
CTA/Managed Futures	18 (11%)	√	5** (27%)	0.00
Distressed Securities	2 (1%)	√	√	0.00
Emerging Market Debt (1%)	1 (1%)	√	-	**0.50**
Emerging Market Equity (Alternative funds)	11 (7%)	√	2** (18%)	0.00
Equity Fundamental Market Neutral	22 (13%)	√	1** (4.5%)	0.00
Equity Statistical Arbitrage	3 (2%)	√	2** (66.7%)	0.00
Fixed Income	3 (2%)	√	√	0.00
Fixed Income Arbitrage	4 (2%)	√	√	0.00
Long Biased Equity	2 (1%)	√	√	0.00
Long/Short Equity	67 (41%)	√	4** (6%)	0.00
Macro	8 (5%)	√	2** (25%)	0.00
Mortgage-Backed Arbitrage	1 (1%)	√	√	0.00
Multi-Style	17 (10%)	√	1** (6%)	1** (6%)
Short Biased Equity	1 (1%)	√	√	0.00
S&P 500	1	√	√	0.00
MSCI Europe	1	√	√	0.00
JP Morgan GABI	1	√	-	0.008

Legend: "√" means all returns of individual hedge funds possess the feature;
"-"means return doesn't possess the feature;
"**" means all returns of the group possess the feature except 1% of observations of the group.

This table represents summary statistics for weekly returns of individual hedge funds and three indexes from 04/01/2008 till 14/12/2012. Percentage into the parenthesis represents the share of hedge funds belonging to the specific management style group. *N* is the number of hedge funds. Excess of kurtosis represents *fat tails*; negative skewness indicates that the *left tail* of the probability density function is longer or fatter than the right side. Jarque-Bera (JB) test is a goodness-of-fit test of whether distribution $\sim \mathcal{N}(m, \sigma^2)$. When JB ***p***-value is below the significance level of 5%, $\mathcal{H}_0$: distribution $\sim \mathcal{N}(m, \sigma^2)$ can be rejected.

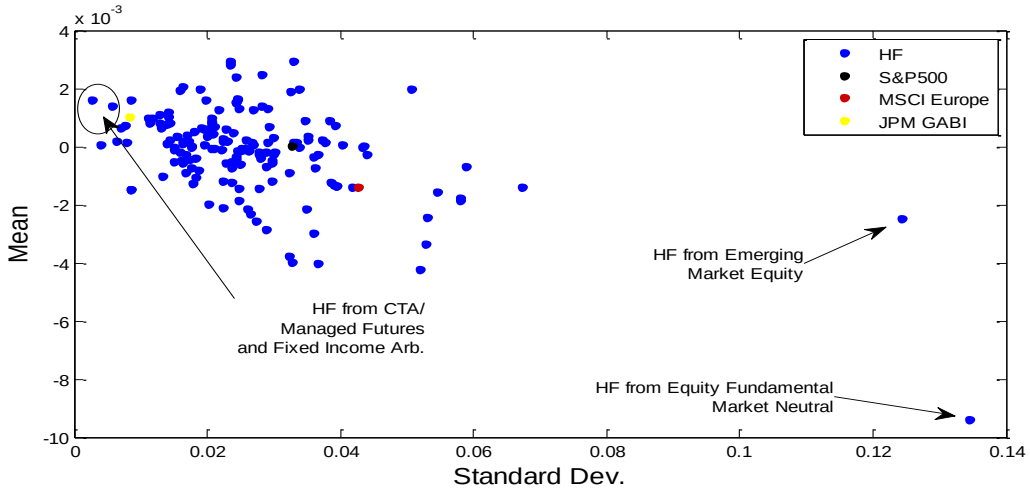


Figure 1.3: Mean and standard deviation of returns of 165 individual hedge funds and three benchmarks
Source: Author's computations.

There is a negative relation between kurtosis and symmetry: the fatter is the tail of the distribution (the higher is the excess of kurtosis) of a return the more the distribution is skewed to the left (Figure 1.4). We notice that there are some individual hedge funds that are much more risky than the two equity and one bond benchmarks. The excess of kurtosis is greater than the benchmark and the distribution is more skewed to the left than the benchmark. The outliers, which underperform the benchmarks, are two hedge funds from Capital Structure/Credit Arbitrage style, three from Multi-Style strategy, one hedge fund form Emerging Market Equity, Fixed Income Arbitrage and Long/Short Equity.

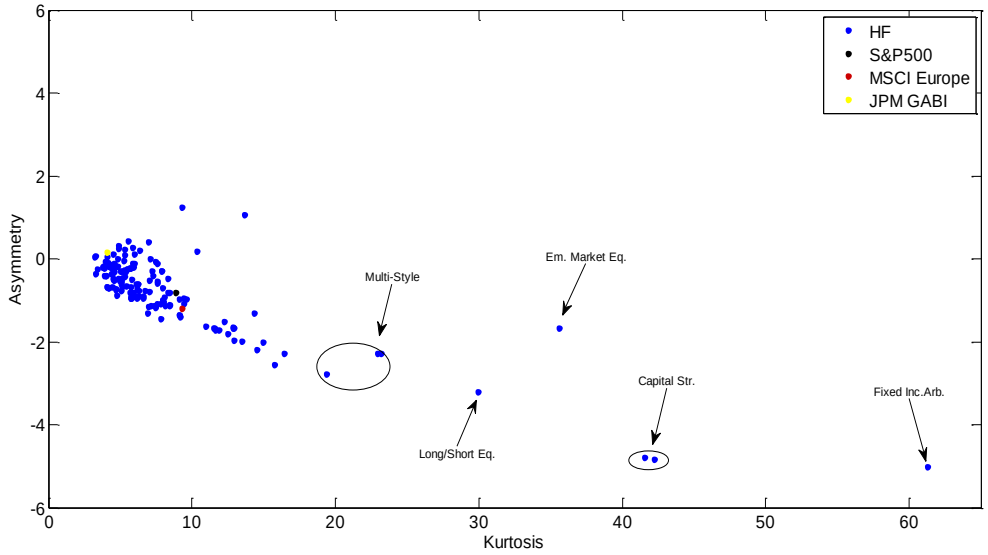


Figure 1.4: Kurtosis and skewness of returns of 165 individual hedge funds and three benchmarks
Source: Author's computations.

In terms of analysis of outliers in terms of standard deviation and third and fourth central moments, the three most risky individual hedge funds are from Emerging Market Equity, Equity Fundamental Market Neutral and Long/Short Equity (see Fig. 1.5). From Fig. 1.5 we also learn that high level of standard deviation corresponds to higher level of kurtosis, there is a positive relationship between standard deviation and kurtosis. We will find a similar relationship on the parameters of the stochastic volatility models.

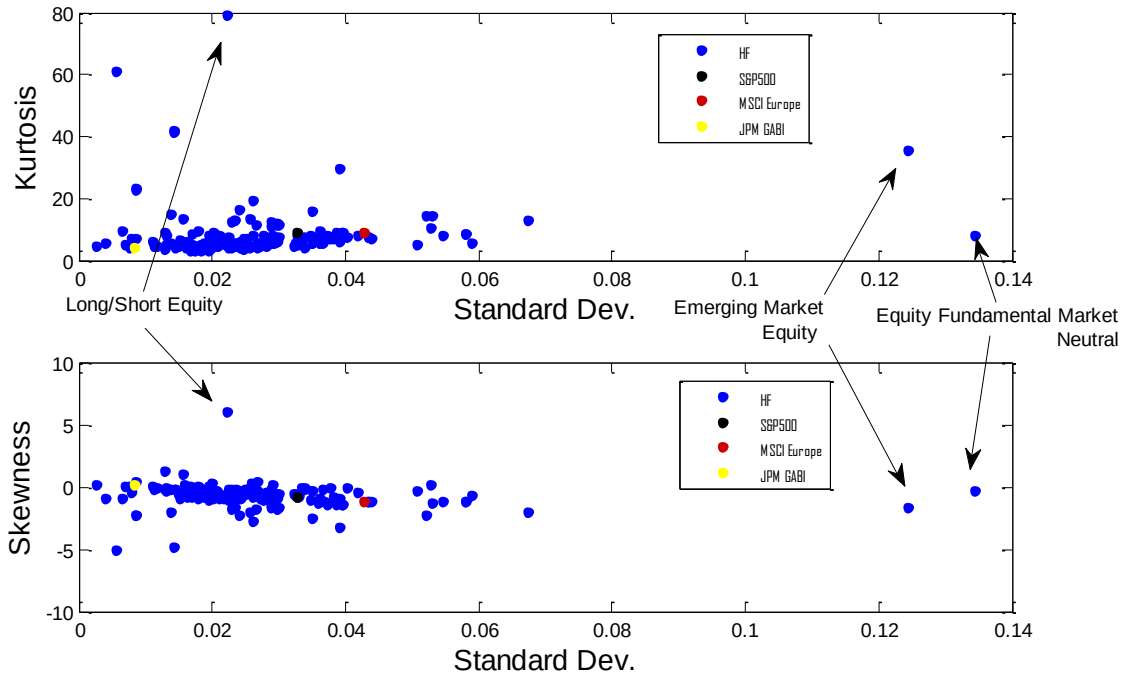


Figure 1.5: Volatility vs. kurtosis and skewness of returns of 165 individual hedge funds and three benchmarks
Source: Author's computations.

Consider now 15 groups of Hedge Funds. Table 1.2 describes mean, standard deviation, skewness, kurtosis and *p*-value of 15 groups of Hedge Funds and 3 market proxies. Categories differ. For instance, Capital Structure/Credit Arbitrage and Fixed Income Arbitrage stand out as having the lowest negative skew (-3.14 and -2.51 respectively) and highest kurtosis (25.25 and 24.99 respectively). The negative skew implies significant losses that are significantly below the average and the excess of kurtosis involves that future returns will be either extremely large or extremely small. In combination, it means that these hedge funds have large risk exposure. Although the two groups of hedge funds exhibit lower standard deviation (2.016% and 1.894%) than the S&P 500, MSCI Europe and JPM GABI indices, they are riskier because of the large excess of kurtosis and a very low skew in

comparison to the benchmarks. A fund with negative skew and excess kurtosis is riskier than a fund with the same standard deviation but smaller negative skew and lower or neutral kurtosis (Tran, V.Q. (2006)).

The two mainly directional strategies are Emerging Market Equity and Short Biased Equity. They show the highest volatility (4.628% and 5.08% respectively) with negative skewness end excess of kurtosis.

Emerging Market Debt exhibits desirable risk characteristics (lower standard deviation than the two equity market benchmarks, that is 1.68%, but higher than the bond market benchmark, positive skewness and very moderate excess of kurtosis (3.24), while recording negative average returns. According to the Jarque-Bera statistics, we can't reject the hypothesis that the distribution is normal.

Equity Statistical Arbitrage strategy in the only one hedge fund strategy showing low risk with relatively low volatility (1.581%), approximately zero skewness (-0.01), but with excess of kurtosis and providing average return higher than S&P 500 and MSCI Europe indices, however still underperforming the bond market.

Distressed Securities management style has the lowest mean: -0.224% and the standard deviation is one of the highest: 3.425%. The lowest standard deviation is reported for the Fixed Income: 1.368% with the mean of -0.027%.

Long/Short Equity, Macro, Mortgage-Backed Arbitrage and Multi-Style classes of hedge funds exhibit quite similar excess of kurtosis and standard deviation in the range from 2.053% to 2.73% thus outperforming the two equity benchmarks. Nevertheless only Mortgage-Backed Arbitrage class of hedge funds among them displays positive return of 0.017%.

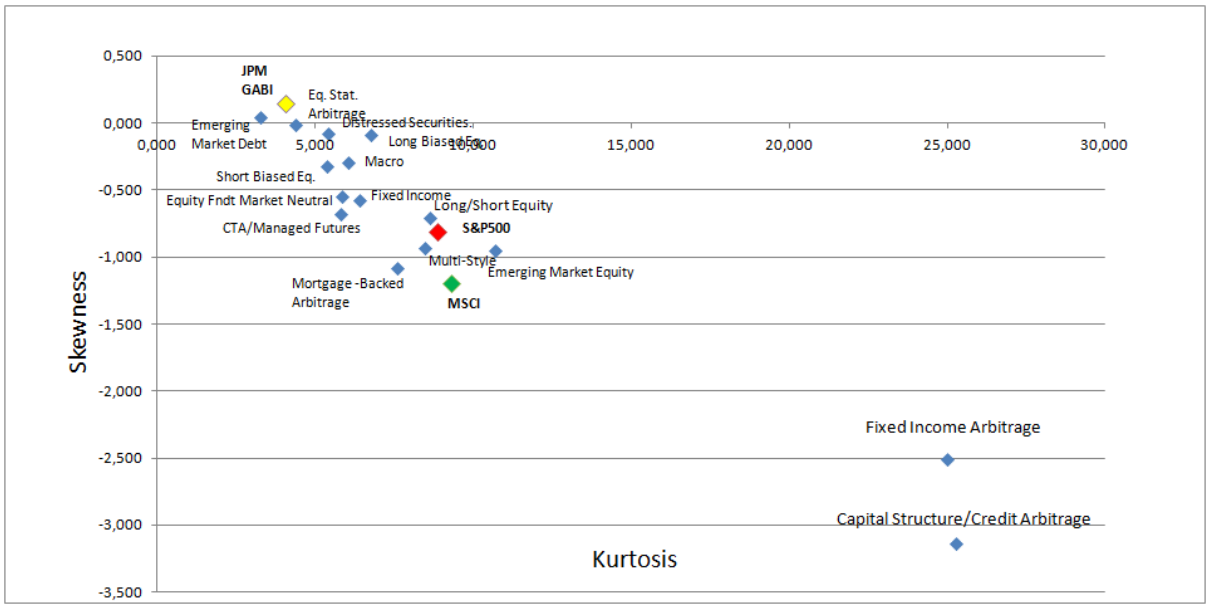
CTA/Managed Futures, Equity Fundamental Market Neutral and Long Biased Equity groups of hedge funds perform positive mean, negative third central moment and excess of kurtosis.

All in all, in comparison with the three market benchmarks, six hedge funds are much more risky than the other funds. Capital Structure/Credit Arbitrage and Fixed Income Arbitrage are risky because of their very low third and very high fourth central moments, and Short Biased Equity, Long Biased Equity, Emerging Market Equity, Distressed Securities classes of hedge funds – because the standard deviation (as a measure of risk of investing in risky asset) of these classes is greater than the standard deviation of the three benchmarks. In comparison with JPM GABI index, no single class of hedge funds can outperform it.

Table 1.2: Summary statistics of groups of hedge fund returns

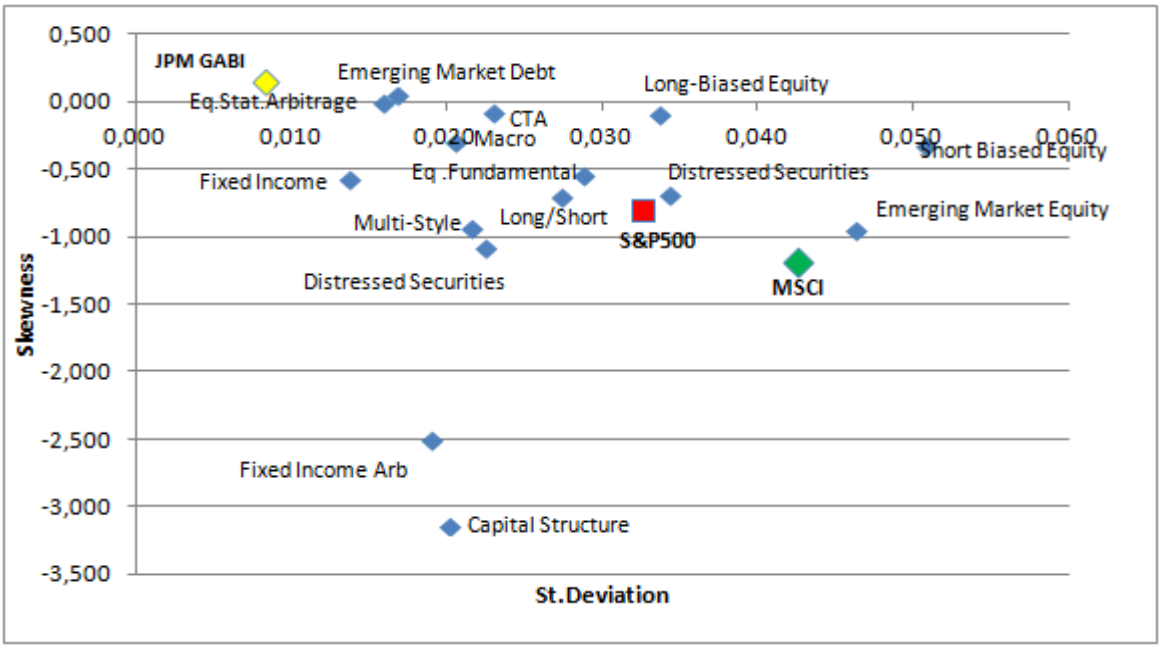
Strategy	st dev (%)	mean (%)	skewness	kurtosis	JB *p*-value
Capital Structure/Credit Arbitrage	2,016	0,064	-3,14	25,25	0,00
CTA/Managed Futures	2,290	0,054	-0,08	5,37	0,00
Distressed Securities	3,425	-0,224	-0,68	5,80	0,00
Emerging Market Debt	1,680	-0,030	0,0444	3,24	0,50
Emerging Market Equity (Alternative funds)	4,628	-0,122	-0,95	10,65	0,00
Equity Fundamental Market Neutral	2,880	0,007	-0,55	5,83	0,00
Equity Statistical Arbitrage	1,581	0,023	-0,01	4,35	0,00
Fixed Income	1,368	-0,027	-0,58	6,38	0,00
Fixed Income Arbitrage	1,894	0,250	-2,51	24,99	0,00
Long Biased Equity	3,361	0,107	-0,09	6,74	0,00
Long/Short Equity	2,730	-0,024	-0,71	8,61	0,00
Macro	2,052	-0,035	-0,30	6,04	0,00
Mortgage-Backed Arbitrage	2,240	0,017	-1,08	7,59	0,00
Multi-Style	2,153	-0,075	-0,93	8,43	0,00
Short Biased Equity	5,080	0,199	-0,32	5,35	0,00
S&P500 Index	3,270	0,001	-0,81	8,89	0,00
MSCI Index	4,270	-0,140	-1,20	9,31	0,00
JPM GABI	0,84	0,1	0,148	4,055	0,008

Figures 1.6, Figure 1.7 and Figure 1.8 give graphical representation of statistical description of the returns of 15 classes of hedge funds.



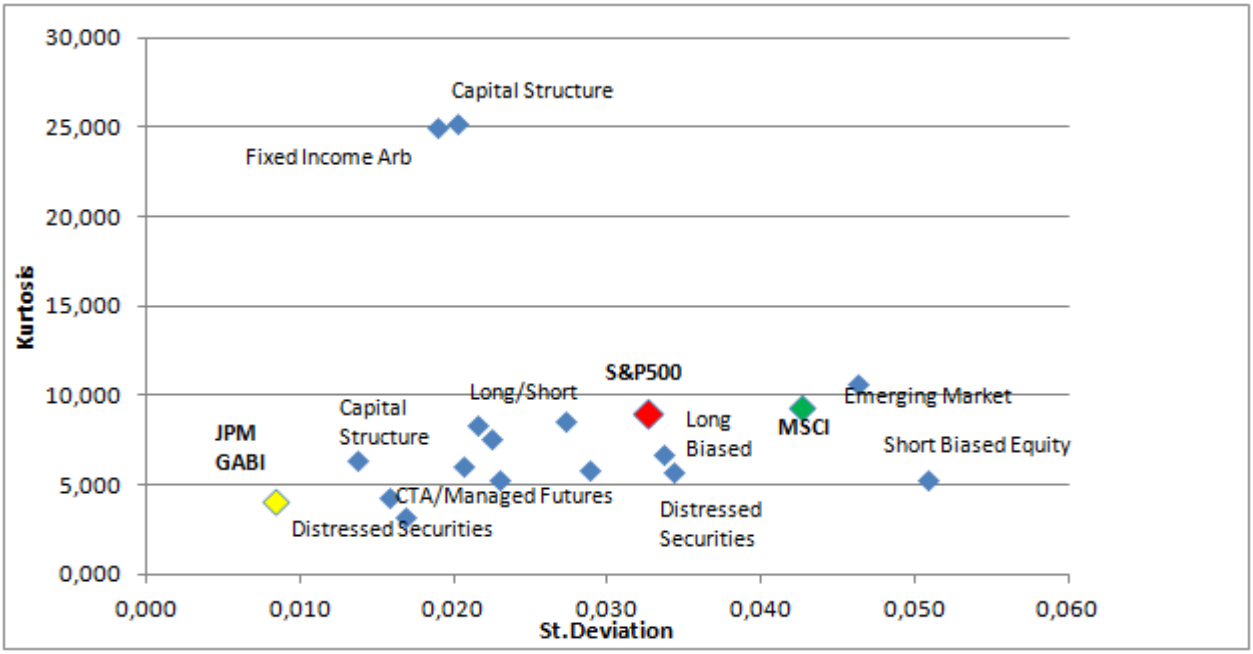
Fixed Income Arbitrage and Capital Structure exhibit lowest skewness and highest excess of kurtosis relatively to the benchmarks. They are considered as to be the riskiest. The probability of high losses of these two classes is very high. The only one class of hedge funds outperforming the three benchmarks is Emerging Market Debt class with almost zero skew and lowest kurtosis.

Figure 1.6: Kurtosis vs. skewness of 15 classes of hedge funds and three market proxies.
Source: Author's computations.



Short Biased Equity and Emerging Market Equity have greater volatility than market proxies. Although Fixed Income Arbitrage and Capital Structure have relatively lower volatility, they exhibit a lower skewness, thus they are risky.

Figure 1.7: Volatility vs. asymmetry of 15 classes of hedge funds and three market proxies
Source: Author's computations.



Emerging Market Equity and Short Biased Equity have higher volatility than the benchmarks. Fixed Income Arbitrage and Capital Structure have highest kurtosis. Thus the four hedge fund classes are risky.

Figure 1.8: Volatility vs. kurtosis of 15 classes of hedge funds and three market proxies
Source: Author's computations.

In this chapter we briefly discussed the history of hedge funds and gave overall review of classification of hedge funds. We have provided and examined the statistical properties of weekly returns of hedge funds and statistical measuring of risk of hedge funds. We have shown that hedge fund returns have non-normal distribution, except one group of Emerging Market Debt. We have also demonstrated that standard deviation which is commonly used to describe the risks of returns is not the one and sufficient condition to display the risk of funds: very relatively low skewness and excess of kurtosis with not relatively low standard deviation can also measure the risk of hedge funds. We will account for some of these stylized facts with the hierarchical multivariate stochastic volatility model presented in the next chapters.

Chapter 2 – Stochastic Volatility

2.1. Introduction

In stochastic volatility (SV) models the mean and the log-volatility equations have separate error terms. However, there is a price to pay in SV model estimation as the evaluation of the likelihood function requires the solution for integral with a dimension equals to the size of the sample.

The intractability of the likelihood function has few reasons: the variance is an unobservable process; the likelihood function is only available in the form of a complex multiple integral. In this context the method of moments (MM) may be inefficient relative to a likelihood-based method of inference because the score function cannot be computed in order to suggest which moment should be used for MM estimation. Also quasi-maximum-likelihood (QML) is no longer reliable because in order to produce a QML estimator during approximation of linear filtering methods it is difficult to evaluate the accuracy of the integral approximation.

Jacquier et al. (1994) found out a solution. They developed new methods for inference and prediction in a simple class of stochastic volatility (SV) models performing Bayesian inference using sampling techniques based upon Markov Chain Monte Carlo (MCMC) simulation methods.

2.2. Bayesian inference

There are two approaches to statistical inference. The first one is the classical approach based on the maximum likelihood principle. A model is estimated by maximizing the likelihood function of the data, and the fitted model is used to make inference. The other approach is Bayesian inference that combines prior belief with data to obtain posterior distributions on which statistical inference is based.

For SV models, the posterior density is $\pi(\theta|h,y)$, where $h = (h_1, \dots, h_T)'$ are the latent volatilities and $y = (y_1, \dots, y_T)'$ are observations, and θ is a vector of all parameters. The posterior density function contains all information in the observations, and provides the basis for the inference and decision making. The Bayes rule factors the posterior distribution as

$$\pi(\theta,h|y) \propto p(y|\theta,h)p(h|\theta)p(\theta),$$

where $p(y|\theta, h)$ is a likelihood function, $p(h|\theta)$ is the conditional distribution of latent process and $p(\theta)$ is the prior distribution of model parameters. The prior distribution allows us to incorporate information into the parameters of interest, prior to the observation y.

2.3. Markov Chain Monte Carlo Simulation Method

The Markov Chain Monte Carlo (MCMC) method is a general simulation method for sampling from posterior distributions and computing posterior quantities of interest. The MCMC methods sample successively from a target distribution. Each sample depends on the previous one, hence the notion of the Markov chain. A Markov chain is a sequence of random variables $\theta_1, \theta_2, \dots, \theta_t$, defined on a state space Θ, such that each element $\theta_t \in \Theta$, depends on all previous elements only through its immediate predecessor θ_{t-1}. A Markov chain is applied to random sampling as a mechanism that traverses randomly through a target distribution without having any memory of where it has been. Where the chain moves next is entirely dependent on its current position on the state space.

Markov chain samples are used in Monte Carlo integration, to approximate an expectation by using empirical averages:

$$\int_{\Theta} g(\theta)p(\theta)\, d\theta \cong \frac{1}{n}\sum_{t=1}^{n} g(\theta_t)$$

where *g(.)* is a function of interest and θ_t , $t = 1, \dots, n$ are samples from *p(θ)* on its support Θ. This approximates the expected value of g(θ). The earliest reference to MCMC simulation occurs in the physics literature. Metropolis and Ulam (1949) and Metropolis et al. (1953) describe what is known as the Metropolis algorithm. The algorithm can be used to generate sequences of samples from the joint distribution of multiple variables, and it is the foundation of MCMC. Hastings (1970) generalized their work, resulting in the Metropolis-Hastings algorithm. Geman and Geman (1984) analyzed image data by using what is now called Gibbs sampling. These MCMC methods first appeared in the mainstream statistical literature in Tanner and Wong (1987).

In Bayesian statistics, there are generally two MCMC algorithms that we use: the Gibbs Sampler and the Metropolis-Hastings algorithm. These methods consist to estimate the posterior distribution and the intractable integrals using simulated samples from the posterior distribution.

2.4. Gibbs Sampling

Following Jacquier et al., many MCMC algorithms have been proposed for different extensions of the SV model. Gibbs sampling (or Gibbs sampler) is the most popular MCMC method; see Geman and Geman (1984) and Gelfand and Smith (1990). In a Gibbs a missing data point can be regarded as a parameter. Similarly, an unobservable variable such as the volatility process of a hedge fund can be regarded as set of parameters.

Denote the three parameters by θ_1, θ_2 and θ_3, let $h = (h_1, \dots, h_T)'$ be the latent volatilities and $y = (y_1, \dots, y_T)'$ the observations. The goal is to approximate the posterior distribution of the parameters so that the fitted model can be used to make inference. Assume three conditional distributions of a single parameter given the others are available. In other words, we assume that the following three distributions are known:

$$f_1\ (\theta_1|\theta_2, \theta_3, y, h)\ ,\ f_2\ (\theta_2|\theta_3, \theta_2, y, h)\ ,\ f_3\ (\theta_3|\theta_1, \theta_2, y, h)\ ,$$

where $f_i\ \left(\theta_i|\ \theta_{j\neq i}, y, h\right)$ denotes the conditional distribution of the parameter θ_i given the other parameters, data and the latent volatilities. We assume to be able to draw a random number from each of the three conditional distributions.

Let $\theta_{2,0}$ and $\theta_{3,0}$ be two arbitrary starting values of θ_2 andθ_3. The Gibbs sampler proceeds as follows:

1. Draw a random sample from$f_1\ \left(\theta_1|\theta_{2,0}, \theta_{3,0}, \boldsymbol{y}, h\right)$. Denote the random draw by$\theta_{1,1}$.
2. Draw a random sample from$f_2\ \left(\theta_2|\theta_{1,1}, \theta_{3,0}, \boldsymbol{y}, h\right)$. Denote the random draw by$\theta_{2,1}$.
3. Draw a random sample from$f_3\ \left(\theta_3|\theta_{1,1}, \theta_{2,1}, \boldsymbol{y}, h\right)$. Denote the random draw by$\theta_{3,1}$.

This completes a Gibbs iteration.

Next, using the new parameters $\theta_{1,1}, \theta_{2,1}$ and $\theta_{3,1}$ as starting values and repeating the prior iteration of random draws, we complete another Gibbs iteration to obtain the updated parameters $\theta_{1,2}, \theta_{2,2}$ and $\theta_{3,2}$. We can repeat the previous iterations for m times to obtain a sequence of m random draws:

$$(\theta_{1,1}, \theta_{2,1}, \theta_{3,1}), \dots, (\theta_{1,m}, \theta_{2,m}, \theta_{3,m}).$$

Under some regularity conditions it can be shown that, for a sufficiently large m, ($\theta_{1,m}$, $\theta_{2,m}$, $\theta_{3,m}$) is approximately equivalent to a random draw from the joint distribution

$f\ (\theta_1, \theta_2, \theta_3 | \ \boldsymbol{y}, h)$ of the three parameters. The regularity conditions are weak; they essentially require that for an arbitrary starting value ($\theta_{1,0}$, $\theta_{2,0}$, $\theta_{3,0}$), the prior Gibbs iterations have a chance to visit the full parameter space. The actual convergence theorem involves using Markov chain theory.

In practice, we use a sufficiently large *n* and discard the first *m* random draws (burn-in sample) of the Gibbs iterations to form a Gibbs sample, for instance,

$$(\theta_{1,m+1},\ \theta_{2,m+1}, \theta_{3,m+1}), \ldots, (\theta_{1,n}, \theta_{2,n}, \theta_{3,n}).$$

Since these realizations form a random sample from the joint posterior distribution $f\ (\theta_1, \theta_2, \theta_3 | \ \boldsymbol{y}, h)$, they can be used to make inference (see Robert and Casella (1999) and Tsay (2005)).

2.5. Metropolis Algorithm

The Gibbs sampler is a special case of a wider class of algorithms called Metropolis algorithms. These algorithms are applicable when the conditional posterior distribution is known up to a normalizing constant; see Metropolis and Ulam (1949) and Metropolis and al. (1953). Suppose we want to draw a random sample from the posterior distribution *f (θ|y),* which contains a complicated normalization constant so that a direct draw is either too time-consuming or infeasible. But there exist an approximate distribution for which random draws are easily available. The Metropolis algorithm generates a sequence of random draws from the approximate distribution whose distributions converge to *f (θ|y).* The algorithm proceeds as follows:

1. Draw a random starting value θ_0 such that $f(\theta_0|y) > 0$.
2. For *t = 1,2,...,*
   a) Draw a candidate sample θ_* from a *known* distribution at iteration *t* given the previous draw θ_{t-1}. Denote the known distribution by $q_t(\theta_t|\theta_{t-1})$, which is called *proposal distribution or jumping distribution* in Gelman et al. (2003).
   b) The proposal distribution is symmetric – that is $q_t(\theta_i|\theta_j) = q_t(\theta_j|\theta_i)$, $\forall\ \theta_i, \theta_j$, *t*. Then the acceptance ratio is defined as

$$r = \frac{f(\theta_*|y)}{f(\theta_{t-1}|y)}.$$

c) Set

$$\theta_t = \begin{cases} \theta_* & \text{with probability } \min(r, 1), \\ \theta_{t-1} & \text{otherwise} \end{cases}$$

Under some regularity conditions, the sequence $\{\theta_t\}$converges in distribution to *f (θ|y);* see Gelman et al. (2003).

Implementation of the algorithm requires calculation of the ratio r for all θ_* and θ_{t-1}, to draw θ_* from the proposal distribution and to draw a random realization from a uniform distribution to determine the acceptance or rejection of θ_*. The normalizing constant of *f (θ|y)* is not needed because only a ratio is used.

The acceptance and rejection rule of the algorithm can be written as follows: i) if the jump from θ_{t-1} to θ_* increases the conditional posterior density, then accept θ_* as θ_t; ii) if the jump decreases the posterior density, then set $\theta_t = \theta_*$ with probability equal to the density ratio r, and set $\theta_t = \theta_{t-1}$otherwise.

Hastings (1970) generalizes the Metropolis algorithm in two ways. First, the proposal distribution does not have to be symmetric. Second, the proposal rule is modified to

$$r = \frac{f(\theta_*|y)/q_t(\theta_*|\theta_{t-1})}{f(\theta_{t-1}|y)/q_t(\theta_{t-1}|\theta_*)} = \frac{f(\theta_*|y)/q_t(\theta_{t-1}|\theta_*)}{f(\theta_{t-1}|y)/q_t(\theta_*|\theta_{t-1})}.$$

This modified Metropolis algorithm is referred to as Metropolis-Hastings algorithm.

2.6. Bayesian Univariate Stochastic Volatility Model

In univariate Stochastic Volatility (SV) asset price dynamics results in the movements of an equity index S_t and its stochastic volatility v_t via continuous time diffusion by a Brownian motion (Lopes and Polson, 2010) :

$$d \log S_t = \mu dt + \sqrt{v_t} dB_t^P \quad (1)$$

$$d \log v_t = \kappa (\gamma - \log v_t)\, dt + \tau dB_t^V \quad (2)$$

Where the parameters governing the volatility evolution are (μ, κ, γ, τ) and the Brownian motions (B_t^P, B_t^V) are possibly correlated.

Data arises in discrete time, thus it is natural to take an Eurler discretization of (1) and (2). Taylor (1987) and West and Harrison (1997) proposed the following non-linear dynamic model:

$$y_t = exp\{h_t/2\}\varepsilon_t \text{ , } \varepsilon_t \sim iid\ N(0, \sigma_\varepsilon^2)$$
$$h_t = \alpha + \varphi h_{t-1} + \eta_t \text{ , } \eta_t \sim iid\ N(0, \sigma_\eta^2)$$

y_t are collection of observed log-returns of a financial asset and h_t is a latent variable, called stochastic log-variance process, $h_0 = 0$ is the initial condition. We assume $\varepsilon_t \perp \eta_s$, $\forall$ *s, t* on the disturbance terms. We use assumption that ε_t and η_t are normally distributed and mutually uncorrelated, because otherwise correlation between the residuals introduces the leverage effect (Nelson, 1991).

This model is popular in financial applications, in describing series with sudden changes in the magnitude of variation of the observed values through a latent linear process (h_t) (Robert and Casella (1999)).

The parameter estimation of SV models was the main obstacle for application of this type of model. Because of the latent volatility, likelihood evaluation requires evaluating an integral with dimension equals to the number of observations.

Given a set of observations $y_1, \dots, y_T$ the likelihood is:

$$L(\alpha, \varphi, \sigma_\eta | y_1, \dots, y_T) = \int_{\mathbb{R}^T} \prod_{t=1}^{T} \frac{1}{\sqrt{(2\pi)\exp(h_t)}} exp\left\{-\frac{1}{2} exp\{-h_t\} y_t^2\right\} \cdot$$
$$\frac{1}{\sqrt{2\pi\sigma_\eta^2}} exp\left\{-\frac{1}{2\sigma_\eta^2}(h_t - \alpha - \varphi h_{t-1})^2\right\} dh_1 \dots dh_T$$

As the integration problem over a large dimensional space arises frequently in latent variable modelling, it can be solved by including the latent variable into the parameter vector. The observed data *y* is augmented by the quantity *h*, latent data, so as to make it easier to calculate likelihood function (it is called *data augmentation* introduced by Tanner and Wong (1987)). Data augmentation principle is used in a "demarginalization" or completion construction, which states: given a probability density f , a density g that satisfies $\int_H g(y,h)dh = f(y)$ is called completion or "demarginalization" of *f*. The density g is chosen so that the full conditionals of g are easy to simulate from and the Gibbs algorithm is implemented on g instead of *f*.

The "demarginalized" complete likelihood function is:

$$L\left(\alpha, \varphi, \sigma_\eta^2 \,|y_1, \ldots, y_T, h_1, \ldots, h_T\right) =$$

$$= \prod_{t=1}^{T} \frac{1}{\sqrt{(2\pi)\exp\left\{\frac{h_t}{2}\right\}}} exp\left\{-\frac{1}{2} exp\{-h_t\} y_t^2\right\} \frac{1}{\sqrt{2\pi\sigma_\eta^2}} exp\left\{-\frac{1}{2\sigma_\eta^2}(h_t - \alpha - \varphi h_{t-1})^2\right\} =$$

$$= \frac{1}{(2\pi)^T \left(\sigma_\eta^2\right)^{T/2}} exp\left\{-\frac{1}{2}\sum_{t=1}^{T}(y_t^2 exp\{-h_t\} + h_t) - \frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(h_t - \varphi - \varphi h_{t-1})^2\right\}$$

where the latent variables (vector of unobserved volatilities) h_1 ,....., h_T are estimated together with the parameters of the SV model.

In the following, we present a Bayesian inference procedure for Univariate SV model. We assume the following *prior* distribution on the parameters α, φ and σ_η:

$$\alpha \sim \mathcal{N}\,(\mu_\alpha; \sigma_\alpha^2) \qquad \textit{(informative prior)}$$

$$\varphi \sim \mathcal{N}\,(\mu_\varphi; \sigma_\varphi^2) \qquad \textit{(informative prior)}$$

$$\sigma_\eta \sim \frac{1}{\sigma_\eta} \qquad \textit{(non-informative improper prior)}$$

where μ_α, σ_α^2 μ_φ, σ_φ^2 and σ_η are known quantities, called *prior hyperparameters*.

In order to simulate samples from the posterior distribution we apply a Gibbs sampling algorithm. Let $\underline{Y}$ = (y_1, ...,y_T)', $\underline{h}$ = (h_1, ..., h_T)' and h_{-t} = (h_1, h_2,....,h_{t-1},h_{t+1},...,h_T), describe the full conditional posterior distributions in the Gibbs sampler and suggest a sampling method for each full conditional.

The *full conditional posterior* distributions of the latent variables are not easy to identify, hence we use a Metropolis-Hastings algorithm to generate samples from these distributions. The resulting algorithm is a Metropolis-Hasting within Gibbs sampling algorithm. The full conditional posterior distributions used in the Gibbs sampling algorithm are given in the following:

(i) Full conditional for α:

$$\pi\,(\alpha \mid \varphi, \sigma_\eta^2\,,\ \underline{Y}, \underline{h}) \propto$$

$$\propto \exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(h_t - \alpha - \varphi h_{t-1})^2\right\} \frac{1}{\sqrt{2\pi}\sigma_\alpha} exp\left\{-\frac{(\alpha-\mu_\alpha)^2}{2\sigma_\alpha^2}\right\} \propto$$

$$\propto \exp\left\{-\frac{1}{2}\left[\alpha^2\left(\frac{T}{\sigma_\eta^2}+\frac{1}{\sigma_\alpha^2}\right)-2\alpha\left(\frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}(h_t-\varphi h_{t-1})+\frac{\mu_\alpha}{\sigma_\alpha^2}\right)\right]\right\}\propto$$

$$\propto \mathcal{N}\left(\overline{\mu_\alpha};\overline{\sigma_\alpha^2}\right),$$

with

$$\overline{\mu_\alpha}=\overline{\sigma_\alpha^2}\left(\frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}(h_t-\varphi h_{t-1})+\frac{\mu_\alpha}{\sigma_\alpha^2}\right),\quad \overline{\sigma_\alpha^2}=\left(\frac{T}{\sigma_\eta^2}+\frac{1}{\sigma_\alpha^2}\right)^{-1}$$

(ii) Full conditional for φ:

$$\pi(\varphi \mid \alpha, \sigma_\eta^2, \underline{Y}, \underline{h})\propto$$

$$\propto exp\left\{-\frac{1}{2\sigma_\eta^2}\sum_{t=1}^{T}(h_t-\alpha-\varphi h_{t-1})^2\right\}\frac{1}{\sqrt{2\pi}\sigma_\varphi}exp\left\{-\frac{1}{2\sigma_\varphi^2}(\varphi-\mu_\varphi)^2\right\}\propto$$

$$\propto exp\left\{-\frac{1}{2}\left[\varphi^2\left(\frac{1}{\sigma_\varphi^2}+\frac{\sum_{t=1}^{T}h_{t-1}^2}{\sigma_\eta^2}\right)-2\varphi\left(\frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}(h_t-\alpha)h_{t-1}\right)+\frac{\mu_\varphi}{\sigma_\varphi^2}\right)\right]\right\}\propto$$

$$\propto \mathcal{N}\left(\overline{\mu_\varphi};\overline{\sigma_\eta^2}\right),$$

with

$$\overline{\mu_\varphi}=\overline{\sigma_\eta^2}\left(\frac{1}{\sigma_\eta^2}\sum_{t=1}^{T}(h_t-\alpha)h_{t-1}\right)+\frac{\mu_\varphi}{\sigma_\varphi^2}\right),\quad \overline{\sigma_\eta^2}=\left(\frac{T}{\sigma_\eta^2}\sum_{t=1}^{T}h_{t-1}^2+\frac{1}{\sigma_\varphi^2}\right)^{-1}$$

(iii) Full conditional for σ_η^2:

$$\pi(\sigma_\eta^2 \mid \alpha,\varphi,\underline{Y},\underline{h})\propto$$

$$\propto \left(\sigma_\eta^2\right)^{-\frac{T}{2}-\frac{1}{2}}exp\left\{-\frac{1}{\sigma_\eta^2}\left(\frac{1}{2}\sum_{t=1}^{T}(h_t-\alpha-\varphi h_{t-1})^2\right)\right\}\propto$$

$$\propto IG\left(\overline{c_{\sigma_\iota}},\overline{d_{\sigma_\iota}}\right)$$

where $IG(.,.)$ indicates the *Inverse Gamma* distribution.

with

$$\overline{c_{\sigma_\iota}}=\frac{T-1}{2},\quad \overline{d_{\sigma_\iota}}=\frac{1}{2}\sum_{t=1}^{T}(h_t-\alpha-\varphi h_{t-1})^2$$

(iv) Full conditional for h_t, t=2,...T-1:

$$\pi(h_t \mid \alpha,\varphi,\sigma_\eta^2\,\underline{Y},\underline{h}_{-t})\propto$$

$$\propto exp\left\{-\frac{h_t}{2}\right\}exp\left\{-\frac{y_t^2}{2}exp\{-h_t\}\right\}exp\left\{-\frac{1}{2\sigma_\eta^2}(h_t-\alpha-\varphi h_{t-1})^2\right\}.$$

$$.\ exp\left\{-\frac{1}{2\sigma_n^2}(h_{t+1}-\alpha-\varphi h_t)^2\right\}\propto$$

$$\propto exp\left\{-\frac{y_t^2}{2}exp\{-h_t\}\right\}.$$

$$. \ exp\left\{-\frac{h_t}{2}-\frac{1}{2\sigma_\eta^2}\left(h_t^2-2h_t(\alpha+\varphi h_{t-1})+h_t^2\varphi^2-2h_t\varphi(h_{t+1}-\alpha)\right)\right\}\propto$$

$$\propto exp\left\{\frac{y_t^2}{2}exp\{-h_t\}\right\}.$$

$$.exp\left\{-\frac{1}{2}\left(h_t^2\frac{1+\varphi^2}{\sigma_\eta^2}-2h_t\frac{1}{\sigma_\eta^2}\left(\alpha(1-\varphi)+\varphi(h_{t-1}+h_{t+1})-\frac{\sigma_\eta^2}{2}\right)\right)\right\}$$

where $\underline{h}_{-t}=(h_1,\dots,h_{t-1},h_{t+1},\dots,h_T)$

We can apply a Metropolis-Hastings step to generate samples from this full conditional distribution. A possible candidate for the proposal distribution *q(h_t)*is a density proportional to the second term in the above expression:

$$q(h_t)\propto \mathcal{N}\left(\mu_{h,t},\sigma_{h,t}^2\right)$$

with

$$\mu_{h,t}=\sigma_{h,t}^2\left(\frac{1}{\sigma_\eta^2}\left(\alpha(1-\varphi)+\varphi(h_{t-1}+h_{t+1})-\frac{\sigma_\eta^2}{2}\right)\right),\quad \sigma_{h,t}^2=\left(\frac{1+\varphi^2}{\sigma_\eta^2}\right)^{-1}$$

In order to find a more efficient proposal distribution we consider second order Taylor approximation of the logarithm of the first factor in the above expression. We consider a Taylor approximation about $\mu_t=\alpha+\varphi h_{t-1}$

$$exp\left\{-\frac{y_t^2}{2}exp\{-h_t\}\right\}\approx$$

$$\approx exp\left\{\frac{y_t^2}{2}\left(exp\{-\mu_t\}-exp\{-\mu_t\}h_t+\frac{1}{2}exp\{-\mu_t\}(h_t-\mu_t)^2\right)\right\}\propto$$

$$\propto exp\left\{-\frac{y_t^2}{2}\left(-exp\{-\mu_t\}h_t+\frac{1}{2}exp\{-\mu_t\}(h_t^2-2\mu_t h_t)\right)\right\}$$

The approximation $\tilde{q}(h_t)$ of the full conditional distribution is

$$\tilde{q}(h_t)\propto$$

$$\propto exp\left\{-\frac{1}{2}\left(h_t^2\left(\frac{1+\varphi^2}{\sigma_\eta^2}+\frac{1}{2}exp\{-\mu_t\}y_t^2\right)\right)\right\}.$$

$$. \ exp\left\{-\frac{1}{2}\left(-2h_t\left(\frac{1}{\sigma_\eta^2}(\alpha(1-\varphi)+\varphi(h_{t-1}+h_{t+1}))\right)\right)\right\}.$$

$$. exp\left\{-\frac{1}{2}\left(-2h_t\left(y_t^2exp\{-\mu_t\}\frac{\mu_t}{2}-\frac{1}{2}\right)\right)\right\}\ \propto$$

$$\propto\ \mathcal{N}\left(\mu_{h,t},\sigma_{h,t}^2\right)$$

with

$$\mu_{h,t} = \sigma_{h,t}^2\left(\frac{1}{\sigma_\eta^2}\left(\alpha(1-\varphi)+\varphi(h_{t-1}+h_{t+1})\right)-\frac{1}{2}+\left(y_t^2 exp\{-\mu_t\}\mu_t/2\right)\right),$$

$$\sigma_{h,t}^2 = \left(\frac{1+\varphi^2}{\sigma_\eta^2}+\frac{1}{2}exp\{-\mu_t\}y_t^2\right)^{-1}$$

(v) Full conditional for h_1

$\pi(h_1|\,\alpha,\varphi,\sigma_\eta^2,\underline{Y},\underline{h}_{-1}) \propto$

$$\propto exp\left\{-\frac{h_1}{2}\right\}exp\left\{-\frac{y_1^2}{2}exp\{-h_1\}\right\}exp\left\{-\frac{1}{2\sigma_\eta^2}(h_1-\alpha)^2\right\}\propto$$

$$\propto exp\left\{-\frac{1}{2\sigma_\eta^2}(h_2-\alpha-\varphi h_1)^2\right\}\propto$$

$$\propto exp\left\{-\frac{y_1^2}{2}exp\{-h_1\}\right\}exp\left\{-\frac{1}{2\sigma_\eta^2}(h_2-\alpha-\varphi h_1)^2-\frac{1}{2\sigma_\eta^2}(h_1-\alpha)^2\right\}\propto$$

$$\propto exp\left\{-\frac{y_1^2}{2}exp\{-h_1\}\right\}exp\left\{-\frac{1}{2\sigma_\eta^2}\left[h_1^2(\varphi^2+1)-2h_1\left(\alpha+\varphi(h_2-\alpha)+\sigma_\eta^2\right)\right]\right\}\propto$$

$$\propto \mathcal{N}\left(\frac{\alpha+\varphi(h_2-\alpha)+\sigma_\eta^2}{1+\varphi^2},\left(\frac{\varphi^2+1}{\sigma_\eta^2}\right)^{-1}\right)$$

(vi) Full conditional for h_T:

$\pi(h_T|\,\alpha,\varphi,\sigma_\eta^2,\underline{Y},\underline{h}_{-T}) \propto$

$$\propto exp\left\{-\frac{h_T}{2}\right\}exp\left\{-\frac{y_T^2}{2}exp\{-h_T\}\right\}exp\left\{-\frac{1}{2\sigma_\eta^2}(h_T-\alpha-\varphi h_{T-1})^2\right\}\propto$$

$$\propto exp\left\{-\frac{h_T}{2}\right\}exp\left\{-\frac{y_T^2}{2}\right\}\mathcal{N}\left(\alpha+\varphi h_{T-1},\sigma_\eta^2\right)$$

2.5. Hierarchical Multivariate Stochastic Volatility Model

We extend the univariate Bayesian SV model given in the previous section to the multivariate setup. We follow a hierarchical modeling approach. Key of the advantage of hierarchical model is that Gibbs sampler is particularly well adapted here. Hierarchical models are stochastic structures in which the distribution *f* can be decomposed as

$$f(y) = \int f_1(y|\theta_1)f_2(\theta_1|\theta_2)\dots f_{\mathbb{I}}(\theta_{\mathbb{I}}|\theta_{\mathbb{I}+1})f_{\mathbb{I}+1}(\theta_{\mathbb{I}+1})d\theta_1\dots d\theta_{\mathbb{I}+1}$$

where θ_j is a parameter and $\mathbb{I}_{\geq 1}$ is the hierarchy level. In our case we use this model to analyze the dependence across assets, by using latent factors in the parameters. The

hierarchical structure leads to a simple Gibbs sampling procedure. Multivariate SV model writes as follows:

$$y_{tj} = exp\{h_{tj}/2\}\varepsilon_{tj}\ ,\ \varepsilon_{tj} \sim iid\ \mathcal{N}(0,\ \sigma_{\varepsilon j}^2)$$
$$h_{tj} = \alpha_j + \varphi_j h_{t-1j} + \eta_{tj}\ ,\ \eta_{tj} \sim iid\ \mathcal{N}(0,\ \sigma_{\eta j}^2)$$

where the index $j = 1, \dots, J$ indicates an hedge fund among J hedge funds in the sample, y_{tj} are fund-specific log-returns, h_{tj} fund-specific log-variances. We assume independence on the disturbance terms: $\varepsilon_{tj} \perp \eta_{sj}$, $\forall$ s,t and $\forall\, j$. We use assumption that ε_{tj} and η_{tj} are normally distributed and mutually uncorrelated, because otherwise correlation between the residuals introduces the leverage effect (Nelson, 1991). Moreover:

1st assumption: $\varepsilon_{tj} \perp \varepsilon_{s,i}\ \forall\ t, s, \forall\ i \neq j$

2nd assumption: $\eta_{tj} \perp \eta_{s,i}\ \forall\ t, s, \forall\ i \neq j$

In the hierarchical specification we assume the following *hierarchical prior* distribution for the parameters α_j, φ_j and $\sigma_{\eta j}$:

$\alpha_j \sim \mathcal{N}\ (\alpha_0 + \underline{\alpha}_j; \sigma_\alpha^2)$ *(informative prior)*

$\varphi_j \sim \mathcal{N}\ (\varphi_0 + \underline{\varphi}_j; \sigma_\varphi^2\ )$ *(informative prior)*

$\sigma_{\eta_j} \sim \frac{1}{\sigma_{\eta_j}}$ *(non-informative improper prior)*

where α_0, σ_α^2 φ_0, σ_φ^2 and σ_{η_j} quantities known as hyperparameters in Bayesian inference.

The *prior distribution* on the common factors is:

$$\alpha_0 \sim\ \mathcal{N}\ (m_0; s_0^2),$$
$$\varphi_0 \sim \mathcal{N}(f_0; v_0^2),$$

with $m_0 = f_0 = 0$ and variances $s_0^2 = v_0^2 = 0.001$.

Using hyperprior distributions to estimate prior distributions is known as *hierarchical Bayes*. In theory, this process could continue further, using hyperprior distributions to estimate the hyperprior distributions. Estimating priors through hyperpriors, and from the data, is a method to elicit the optimal prior distributions.

The structure of the model is the following:

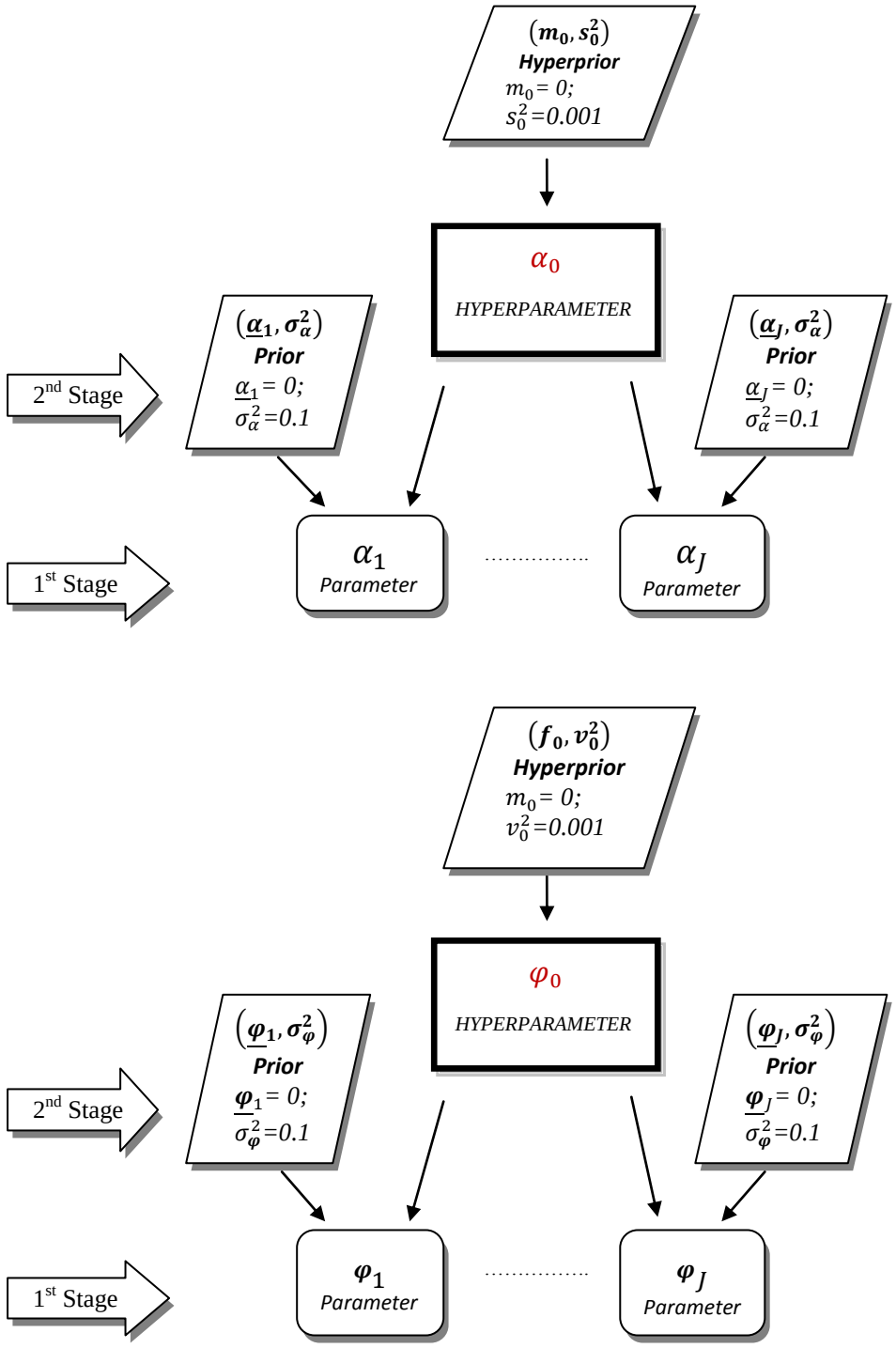


Figure 2.1: Representation of Hierarchical Structure and setting of the prior parameters

The complete likelihood function of the panel of hedge funds is:

$$L\left(\alpha_j, \varphi_j, \sigma_{\eta_j}^2 \middle| y_{11}, \dots, y_{TJ}, h_{11}, \dots, h_{TJ}\right) =$$

$$= \prod_{j=1}^{J}\prod_{t=1}^{T} \frac{1}{\sqrt{(2\pi)\,exp(h_{tj})}} exp\left\{-\frac{1}{2}exp\{-h_{tj}\}y_{tj}^2\right\} \frac{1}{\sqrt{2\pi\sigma_{\eta_j}^2}} exp\left\{-\frac{1}{2\sigma_{\eta_j}^2}\left(h_{tj} - \alpha_j - \varphi_j h_{t-1,j}\right)^2\right\} =$$

$$= \frac{1}{(2\pi)^T\left(\sigma_{\eta_j}^2\right)^{T/2}} exp\left\{-\frac{1}{2}\sum_{j=1}^{J}\sum_{t=1}^{T}\left(y_{tj}^2 exp\{-h_{tj}\} + h_{tj}\right) - \frac{1}{2\sigma_{\eta j}^2}\sum_{j=1}^{J}\sum_{t=1}^{T}\left(h_{tj} - \varphi_j - \varphi_j h_{t-1,j}\right)^2\right\}$$

Inference on the posterior distribution relies upon Monte Carlo approximation methods. More specifically we apply a Gibbs sampling with the following full conditional distributions:

(i) Full conditional for $\alpha_j, j = 1, \dots, J$:

$$\pi\,(\alpha_j \mid \varphi_j, \alpha_0, \underline{\alpha}_j, \sigma_{\eta_j}^2, \underline{Y}, \underline{h}) \propto$$

$$\propto exp\left\{-\frac{1}{2\sigma_{\eta j}^2}\sum_{t=1}^{T}\left(h_{t,j} - \alpha_j - \varphi_j h_{t-1,j}\right)^2\right\}\frac{1}{\sqrt{2\pi}\sigma_{\alpha_j}} exp\left\{-\frac{\left(\alpha_j - \alpha_0 - \underline{\alpha}_j\right)^2}{2\sigma_{\alpha_j}^2}\right\} \propto$$

$$\propto exp\left\{-\frac{1}{2}\left[\alpha_j^2\left(\frac{T}{\sigma_{\eta_j}^2} + \frac{1}{\sigma_{\alpha_j}^2}\right) - 2\alpha_j^2\left(\frac{1}{\sigma_{\eta_j}^2}\sum_{t=1}^{T}\left(h_{t,j} - \varphi_j h_{t-1,j}\right) + \frac{\alpha_0 + \underline{\alpha}_j}{\sigma_{\alpha}^2}\right)\right]\right\} \propto$$

$$\propto \mathcal{N}\,(\overline{\mu_j}; \overline{\sigma_{\alpha_j}^2}),$$

with

$$\overline{\mu_j} = \overline{\sigma_{\alpha_j}^2}\left(\frac{1}{\sigma_{\eta_j}^2}\sum_{t=1}^{T}\left(h_{t,j} - \varphi_j h_{t-1,j}\right) + \frac{\alpha_0 + \underline{\alpha}_j}{\sigma_{\alpha_j}^2}\right), \quad \overline{\sigma_{\alpha j}^2} = \left(\frac{T}{\sigma_{\eta_j}^2} + \frac{1}{\sigma_{\alpha_j}^2}\right)^{-1}$$

(ii) Full conditional for $\varphi_j, j = 1, \dots, J$:

$$\pi\,(\varphi_j \mid \alpha_j, \varphi_0, \sigma_{\eta_j}^2, \underline{\varphi}_j, , \underline{Y}, \underline{h}) \propto$$

$$\propto \exp\left\{-\frac{1}{2\sigma_{\eta_j}^2}\sum_{t=1}^{T}\left(h_{t,j} - \alpha_j - \varphi_j h_{t-1,j}\right)^2\right\}\frac{1}{\sqrt{2\pi}\sigma_{\varphi_j}} exp\left\{-\frac{1}{2\sigma_{\varphi_j}^2}\left(\varphi_j - \varphi_0 - \underline{\varphi}_j\right)^2\right\} \propto$$

$$\propto \exp\left\{-\frac{1}{2}\left[\varphi_j^2\left(\frac{1}{\sigma_{\varphi_j}^2} + \frac{\sum_{t=1}^{T} h_{t-1}^2}{\sigma_{\eta_j}^2}\right) - 2\varphi_j^2\left(\frac{1}{\sigma_{\eta_j}^2}\sum_{t=1}^{T}\left(h_{t,j} - \alpha_j\right)h_{t-1,j}\right) + \frac{\varphi_0 + \underline{\varphi}_j}{\sigma_{\varphi}^2}\right)\right]\right\} \propto$$

$$\propto \mathcal{N}\,(\overline{\mu_j}; \overline{\sigma_{\eta j}^2}),$$

with

$$\overline{\mu_j} = \overline{\sigma_{\eta_j}^2}\left(\frac{1}{\sigma_{\eta_j}^2}\sum_{t=1}^{T}(h_{t,j}-\alpha_j)h_{t-1,j}\right)+\frac{\varphi_0+\varphi_j}{\sigma_\varphi^2}\Big), \quad \overline{\sigma_{\eta_j}^2} = \left(\frac{T}{\sigma_{\eta_j}^2}\sum_{t=1}^{T}h_{t-1,j}^2+\frac{1}{\sigma_\varphi^2}\right)^{-1}$$

(iii) Full conditional for $\sigma_{\eta_j}^2, j = 1,\dots,J$:

$$\pi\left(\sigma_{\eta_j}^2 \middle| \alpha_j,\varphi_j,\alpha_0,\varphi_0,\underline{Y},\underline{h}\right) \propto$$

$$\propto \prod_{t=1}^{T}\frac{1}{2\sigma_{\eta j}^2}exp\left\{-\frac{1}{2\sigma_{\eta j}^2}\left(h_{t,j}-\alpha_j-\varphi_j h_{t-1,j}\right)^2\right\}$$

$$\propto \left(\sigma_{\eta_j}^2\right)^{-\frac{T}{2}-\frac{1}{2}}exp\left\{-\frac{1}{\sigma_{\eta j}^2}\left(\frac{1}{2}\sum_{t=1}^{T}\left(h_{t,j}-\alpha_j-\varphi_j h_{t-1,j}\right)^2\right)\right\}$$

$$\propto IG\left(\overline{c_{\sigma_\iota}},\overline{d_{\sigma_\iota}}\right)$$

with

$$\overline{c_{\sigma_\iota}} = \frac{T-1}{2}, \quad \overline{d_{\sigma_\iota}} = \frac{1}{2}\sum_{t=1}^{T}\left(h_{t,j}-\alpha_j-\varphi_j h_{t-1,j}\right)^2$$

(iv) Full conditional for h_{tj}, *t=2,...T-1*, j = *1,...,J:*

$$\pi\left(h_{t,j} \middle| \alpha_j,\varphi_j,\alpha_0,\varphi_0,\underline{Y},\underline{h_{-t,j}}\right) \propto$$

$$\propto exp\left\{-\frac{h_{t,j}}{2}\right\}exp\left\{-\frac{y_{t,j}^2}{2}exp\{-h_{t,j}\}\right\}exp\left\{-\frac{1}{2\sigma_\eta^2}\left(h_{t,j}-\alpha_j-\varphi_j h_{t-1,j}\right)^2\right\}.$$

$$.\ exp\left\{-\frac{1}{2\sigma_{nj}^2}\left(h_{t+1,j}-\alpha_j-\varphi_j h_{t,j}\right)^2\right\} \propto$$

$$\propto exp\left\{-\frac{y_{t,j}^2}{2}exp\{-h_{t,j}\}\right\}.$$

$$.\ exp\left\{-\frac{h_{t,j}}{2}-\frac{1}{2\sigma_{\eta j}^2}\left(h_{t,j}^2-2h_{t,j}\left(\alpha_j+\varphi_j h_{t-1,j}\right)+h_{t,j}^2\varphi^2-2h_{t,j}\varphi_j\left(h_{t+1,j}-\alpha_j\right)\right)\right\}$$

$$\propto exp\left\{\frac{y_{t,j}^2}{2}exp\{-h_{t,j}\}\right\}$$

$$.\ exp\left\{-\frac{1}{2}\left(h_{t,j}^2\frac{1+\varphi_j^2}{\sigma_{\eta j}^2}-2h_{t,j}\frac{1}{\sigma_{\eta j}^2}\left(\alpha_j(1-\varphi_j)+\varphi_j\left(h_{t-1,j}+h_{t+1,j}\right)-\frac{\sigma_{\eta j}^2}{2}\right)\right)\right\}$$

where $\underline{h_{-t,j}} = \left(h_{1,j},\dots,h_{t-1,j},h_{t+1,j},\dots,h_{T,J}\right)$

Applying Metropolis-Hastings steps, we get a density proposal distribution *q($h_{t,j}$):*

$$q(h_{t,j}) \propto \mathcal{N}\left(\mu_{h,j,t}, \sigma^2_{h,j,t}\right)$$

with

$$\mu_{h,j,t} = \sigma^2_{h,j,t}\left(\frac{1}{\sigma^2_{\eta,j}}\left(\alpha_j(1-\varphi_j) + \varphi_j\left(h_{t-1,j} + h_{t+1,j}\right) - \frac{\sigma^2_{\eta j}}{2}\right)\right), \quad \sigma^2_{h,j,t} = \left(\frac{1+\varphi_j^2}{\sigma^2_{\eta j}}\right)^{-1}$$

We consider second order Taylor approximation of the logarithm of the first factor in the above expression. We consider a Taylor approximation about $\mu_{t,j} = \alpha_j + \varphi_j h_{t-1,j}$

$$exp\left\{-\frac{y^2_{t,j}}{2} exp\{-h_{t,j}\}\right\} \approx$$

$$\approx exp\left\{\frac{y^2_{t,j}}{2}\left(exp\left\{-\mu_{t,j}\right\} - exp\left\{-\mu_{t,j}\right\} h_{t,j} + \frac{1}{2} exp\left\{-\mu_{t,j}\right\}\left(h_{t,j} - \mu_{t,j}\right)^2\right)\right\} \propto$$

$$\propto exp\left\{-\frac{y^2_{t,j}}{2}\left(-exp\left\{-\mu_{t,j}\right\} h_{t,j} + \frac{1}{2} exp\left\{-\mu_{t,j}\right\}\left(h^2_{t,j} - 2\mu_{t,j} h_{t,j}\right)\right)\right\}$$

The approximation $\tilde{q}(h_{t,j})$ of the full conditional distribution is

$$\tilde{q}(h_{t,j}) \propto \mathcal{N}\left(\mu_{h,j,t}, \sigma^2_{h,j,t}\right)$$

with

$$\mu_{h,j,t} = \sigma^2_{h,j,t}\left(\frac{1}{\sigma^2_{\eta j}}\left(\alpha_j(1-\varphi_j) + \varphi_j\left(h_{t-1,j} + h_{t+1,j}\right)\right) - \frac{1}{2} + \left(y^2_{t,j} exp\left\{-\mu_{t,j}\right\}\mu_{t,j}/2\right)\right),$$

$$\sigma^2_{h,j,t} = \left(\frac{1+\varphi_j^2}{\sigma^2_{\eta j}} + \frac{1}{2} exp\left\{-\mu_{t,j}\right\} y^2_{t,j}\right)^{-1}$$

(v) Full conditional for h_1, $j = 1,\ldots,J$:

$$\pi\left(h_{1,j} \mid \alpha_j, \varphi_j, \sigma^2_{\eta j}, \underline{Y}, \underline{h}_{-1,j}\right) \propto$$

$$\propto exp\left\{-\frac{h_{1,j}}{2}\right\} exp\left\{-\frac{y^2_{1,j}}{2} exp\{-h_{1,j}\}\right\} exp\left\{-\frac{1}{2\sigma^2_{\eta j}}\left(h_{1,j} - \alpha_j\right)^2\right\}.$$

$$. \; exp\left\{-\frac{1}{2\sigma^2_{\eta j}}\left(h_{2,j} - \alpha_j - \varphi_j h_{1,j}\right)^2\right\} \propto$$

$$\propto \mathcal{N}\left(\frac{\alpha_j + \varphi_j(h_{2,j} - \alpha_j) + \sigma^2_{\eta j}}{1+\varphi^2}, \left(\frac{\varphi_j^2+1}{\sigma^2_{\eta j}}\right)^{-1}\right)$$

(vi) Full conditional for h_T, $j = 1,\ldots,J$:

$$\pi\left(h_T \mid \alpha_j, \varphi_j, \sigma^2_{\eta j,} \; \underline{Y}, \underline{h}_{-T,J}\right) \propto$$

$$\propto \; exp\left\{-\frac{h_{T,j}}{2}\right\} exp\left\{-\frac{y^2_{T,j}}{2} exp\{-h_{T,j}\}\right\} exp\left\{-\frac{1}{2\sigma^2_{\eta j}}\left(h_{T,j} - \alpha_j - \varphi_j h_{T-1,j}\right)^2\right\} \propto$$

$$\propto \quad exp\left\{-\frac{h_{T,j}}{2}\right\}exp\left\{-\frac{y_{T,j}^2}{2}\right\}\mathcal{N}\left(\alpha_j + \varphi_j h_{T-1,j}\,, \sigma_{\eta j}^2\right)$$

(vii) Full conditional for α_0:

$$\pi\,(\alpha_0|\,\alpha_1,\ldots,\,\alpha_J) \propto$$

$$\propto exp\left\{-\frac{1}{2\sigma_\alpha^2}\sum_{j=1}^{J}\left(\alpha_j - \underline{\alpha}_j - \alpha_0\right)^2\right\}exp\left\{-\frac{1}{2s_0^2}(\alpha_0 - m_0)^2\right\}\frac{1}{\sqrt{2\pi}s_0} \propto$$

$$\propto exp\left\{-\frac{1}{2}\left[\alpha_0^2\left(\frac{J}{\sigma_\alpha^2} + \frac{1}{s_0^2}\right) - 2\alpha_0\left(\frac{1}{\sigma_\alpha^2}\sum_{j=1}^{J}\left(\alpha_j - \underline{\alpha}_j\right) + \frac{m_0}{s_0^2}\right)\right]\right\} \propto$$

$$\propto \mathcal{N}\left(\overline{m_0}, \overline{s_0^2}\right)$$

with

$$\overline{m_0} = \overline{s_0^2}\left(\frac{1}{\sigma_\alpha^2}\sum_{j=1}^{J}\left(\alpha_j - \underline{\alpha}_j\right) + \frac{m_0}{s_0^2}\right), \quad \overline{s_0^2} = \left(\frac{J}{\sigma_\alpha^2} + \frac{1}{s_0^2}\right)^{-1}$$

(viii) Full conditional for φ_0:

$$\pi\,(\varphi_0|\,\varphi_1,\ldots,\,\varphi_J) \propto$$

$$\propto exp\left\{-\frac{1}{2\sigma_\varphi^2}\left(\varphi_j - \underline{\varphi}_j - \varphi_0\right)^2\right\}\frac{1}{\sqrt{2\pi}v_0}exp\left\{-\frac{1}{2v_0^2}(\varphi_0 - f_0)^2\right\} \propto$$

$$\propto exp\left\{-\frac{1}{2}\left[\varphi_0^2\left(\frac{J}{\sigma_\varphi^2} + \frac{1}{v_0^2}\right) - 2\varphi_0\left(\frac{1}{\sigma_\varphi^2}\sum_{j=1}^{J}\left(\varphi_j - \underline{\varphi}_j\right) + \frac{f_0}{v_0^2}\right)\right]\right\} \propto$$

$$\propto \mathcal{N}\left(\overline{f_0}, \overline{v_0^2}\right)$$

with

$$\overline{f_0} = \overline{v_0^2}\left(\frac{1}{\sigma_\varphi^2}\sum_{j=1}^{J}\left(\varphi_j - \underline{\varphi}_j\right) + \frac{f_0}{v_0^2}\right), \quad \overline{v_0^2} = \left(\frac{J}{\sigma_\varphi^2} + \frac{1}{v_0^2}\right)^{-1}$$

Chapter 3 - Simulation and Application to the Hedge Funds

3.1. Simulation Results

In this part we study, through some simulation experiments, the efficiency of the MCMC algorithm.

In the simulation experiments we proceed as follows. We consider 165 units (e.g., hedge funds) and generate n = 250 observations for each unit from our multivariate SV model. On the synthetic data set we iterate the MCMC algorithm for 300 times. We consider a dataset of 165 units and of 250 observations because this is the typical sample size of the applications to real data we will introduce in the next section. As for the values of the parameters we consider α = - 0.8, φ = 0.6, and σ_η^2 = 0.363 for each unit. These are the true arbitrary chosen parameters and we simulate the model (see Fig. 1).

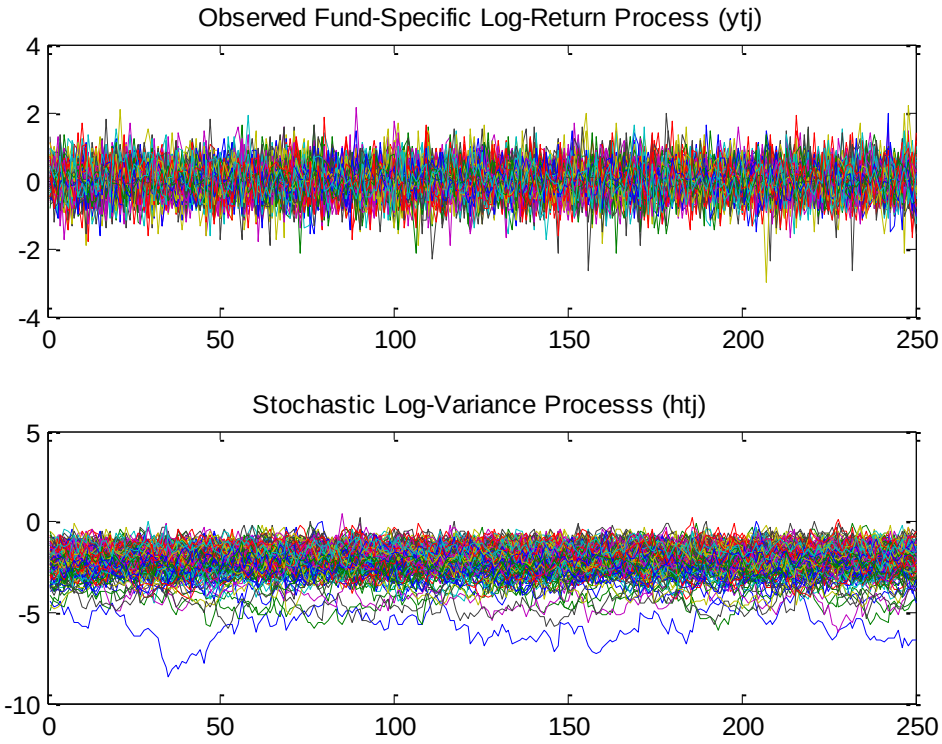


Figure 3.1: Simulated trajectories of 250 observations from the observable process $y_{t,j}$, with j = 1,..., 165, number of hedge funds, with time varying volatility and of the stochastic volatility process $h_{t,j}$. We simulate the model by setting α = - 0.8, φ = 0.6, and σ_η^2 = 0.363 for each hedge fund.

We apply our Gibbs sampler to the synthetic data set represented in Fig. 3.1. We use a burn-in sample of 50 observations to reduce the dependence on the initial conditions of the sampler. Figure 3.2 shows raw output of one independent randomly selected hedge fund of 300 MCMC Gibbs iterations.

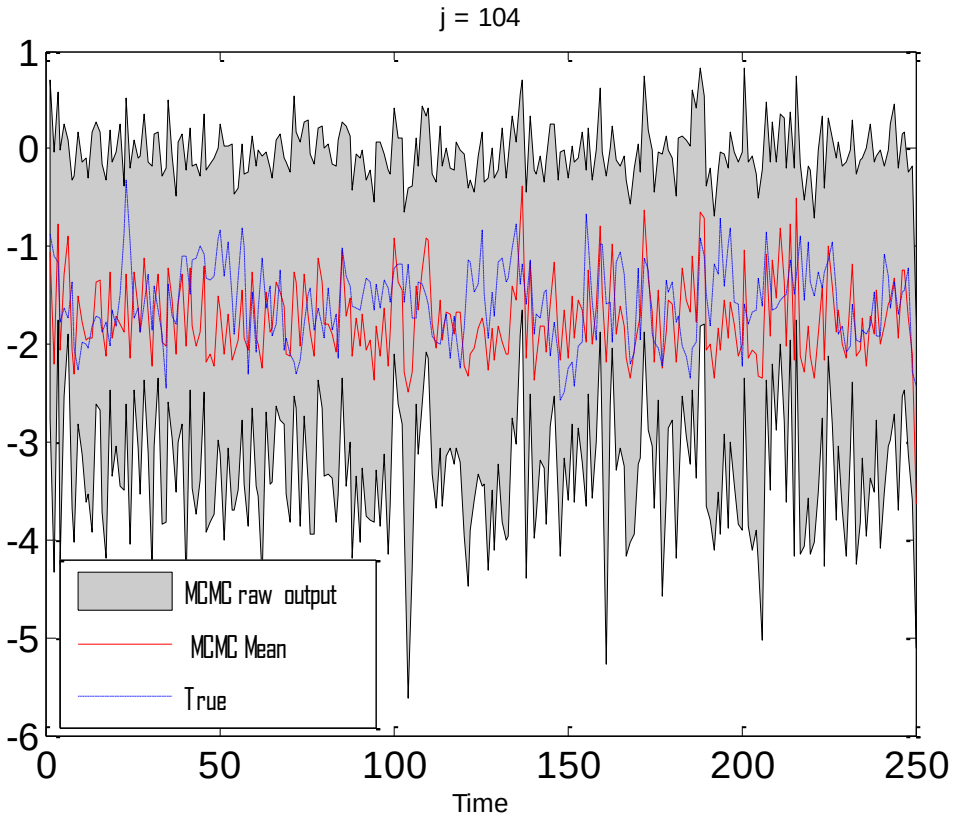


Figure 3.2: MCMC (Gibbs sampler) raw output for the first latent variables $h_{t,104}$, calculated on a synthetic dataset of T = 250 and randomly selected hedge fund number 104. Gray area: 95% high probability density region, blue dashed line: true latent volatility value, and red solid line: posterior mean (estimated latent volatility value).

The true value of the log-volatility process belongs to the 95% high probability density region (Fig. 3.2). Thus there is strong evidence in favour of the hypothesis that the red line (the estimated value) is equal to the blue line (the true value). The estimated parameters $\hat{\alpha}_j$ and $\hat{\varphi}_j$ are represented in Figure 3.3. The estimated parameters (red dots) are very close to the true parameters (blue dots), which shows the effectiveness of our estimation procedure.

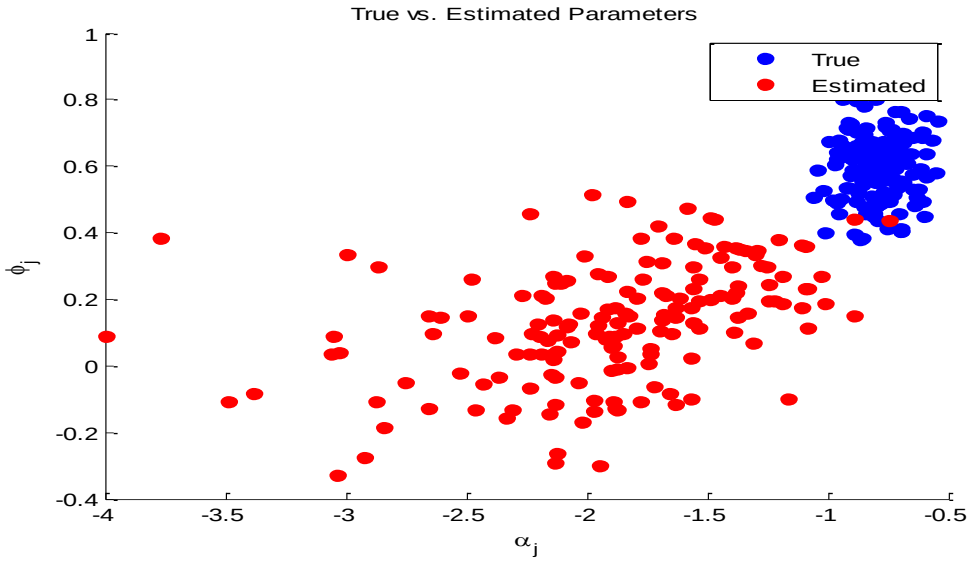


Figure 3.3: True parameters α_j , φ_j and the estimated parameters $\hat{\alpha}_j$, $\hat{\varphi}_j$

3.2. Empirical Application to Hedge Fund Returns

Here we consider weekly observations from 04/01/2008 to 14/12/2012 of the 165 hedge fund returns. As shown in Figure 3.4 (up), the hedge funds log-returns exhibit time variation in the conditional velocity.

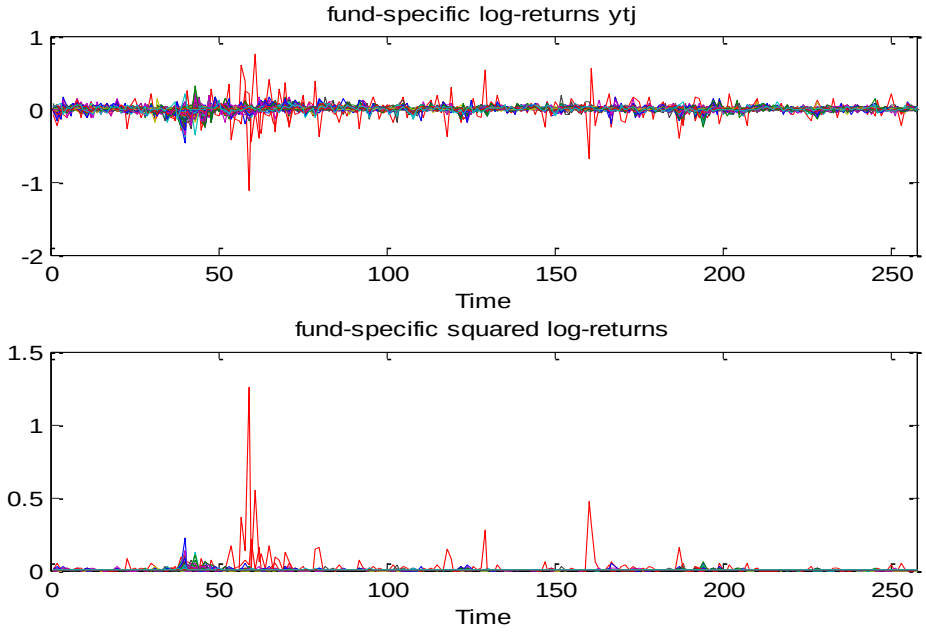


Figure 3.4: Up: Log-return of 165 hedge funds at the weekly frequency for the period 04/01/2008 to 14/12/20012. Bottom: Square of log-return of 165 hedge funds

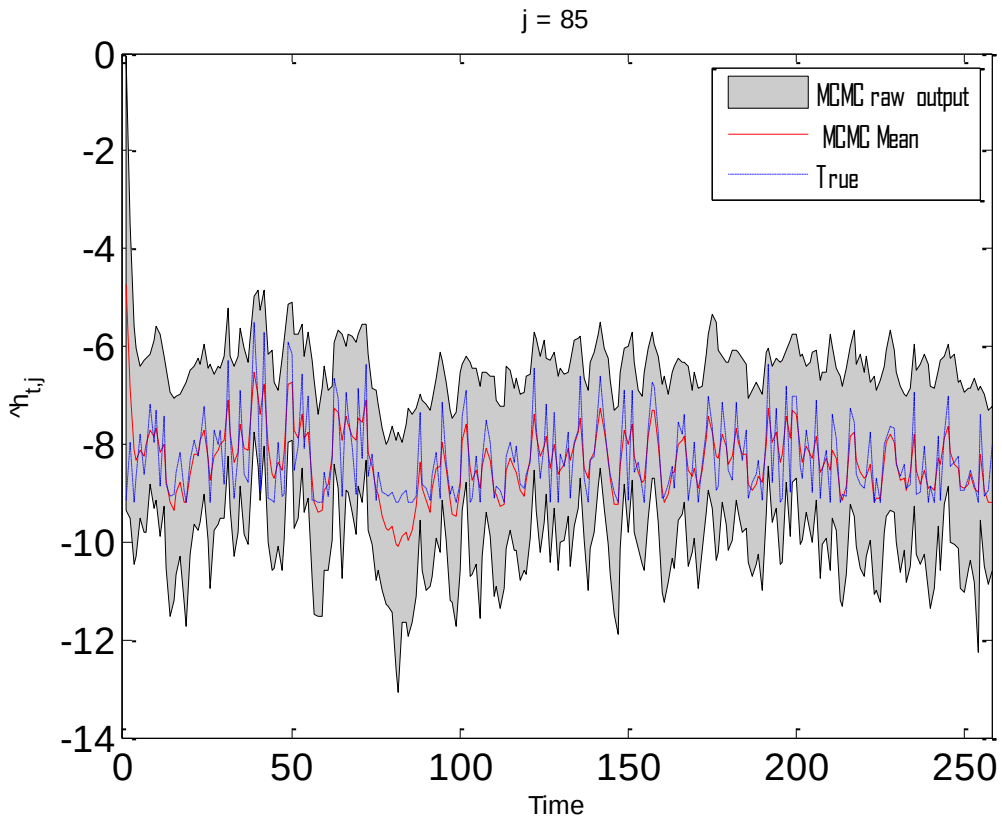


Figure 3.5: MCMC (Gibbs sampler) raw output for the latent variables $\hat{h}_{t,85}$, calculated on a dataset of T = 250 and randomly selected hedge fund number 85. Gray area: 95% high probability density region, blue dashed line: true latent volatility value, and red solid line: posterior mean (estimated latent volatility value).

Figure 3.5 shows the initial value of the log-volatility process (blue line) and the estimated log-volatility process (red line) together with the 95% high probability density region. We can conclude that the rolling estimation of the volatility is a good proxy of the estimated one.

Figure 3.6 graphs the output of 300 MCMC (Gibbs) iterations for the parameters $\alpha_{t,j}$ and $\varphi_{t,j}$. From a graphical inspection, it seems that parameters of each hedge fund stabilize quickly after 200 iterations.

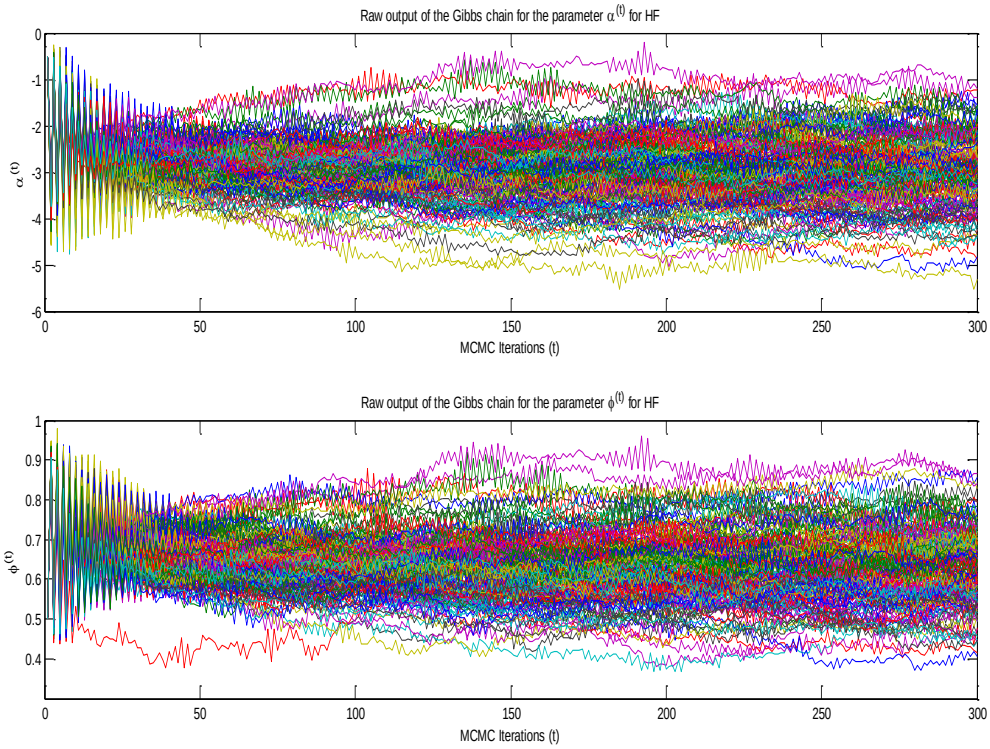


Figure 3.6: Raw MCMC outputs over 300 Gibbs iterations for the parameters $\alpha_{t,j}$ and $\varphi_{t,j}$.

We found that there is a positive relationship between estimated volatility level and estimated persistence of volatility (Fig. 3.7 and Fig. 3.8). The higher the level of volatility, the higher is the persistence. The comparative analysis of outliers leads us to a conclusion that higher value of volatility persistence $\hat{\varphi}_{t,j}$ corresponds to the higher level of kurtosis, calculated in the chapter 1 (Fig. 3.9 a) and b)). We remark that two hedge funds from Capital Structure group, one from Fixed Income Arbitrage exhibit high level of kurtosis and high level of volatility persistence. The hedge funds that exhibit the highest persistence of volatility are mostly from Long/Short Equity and Emerging Market Equity management styles.

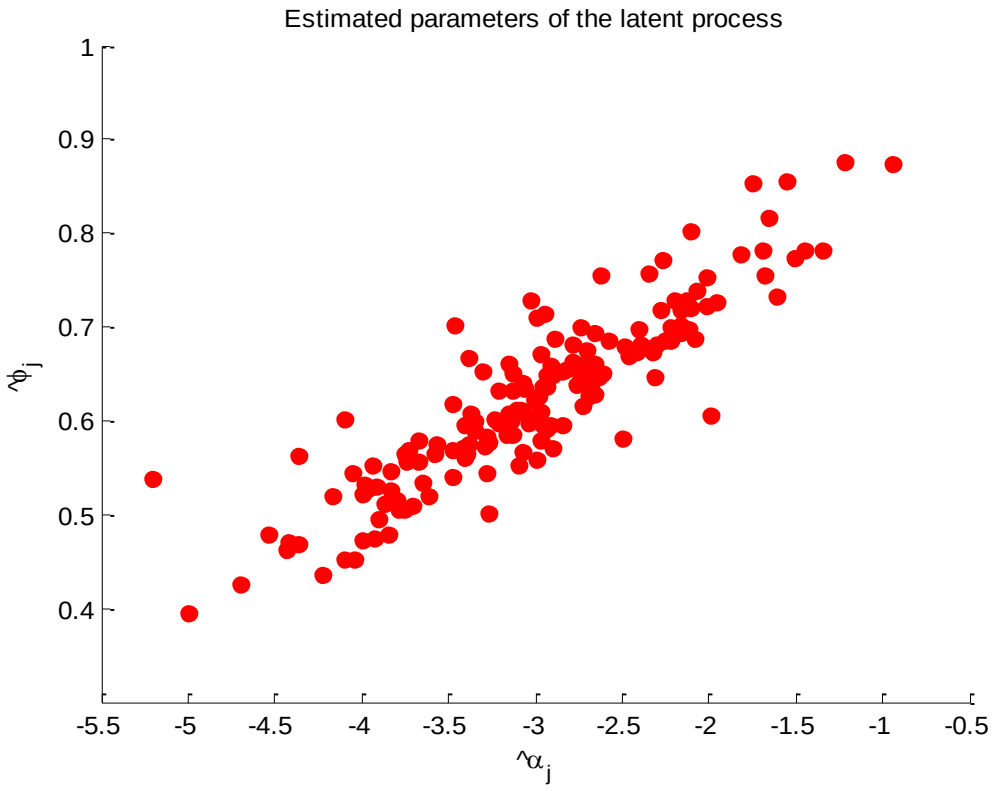


Figure 3.7: True and estimated values of intercept (volatility level) $\hat{\alpha}_{t,j}$ and persistence of volatility $\hat{\varphi}_{t,j}$.

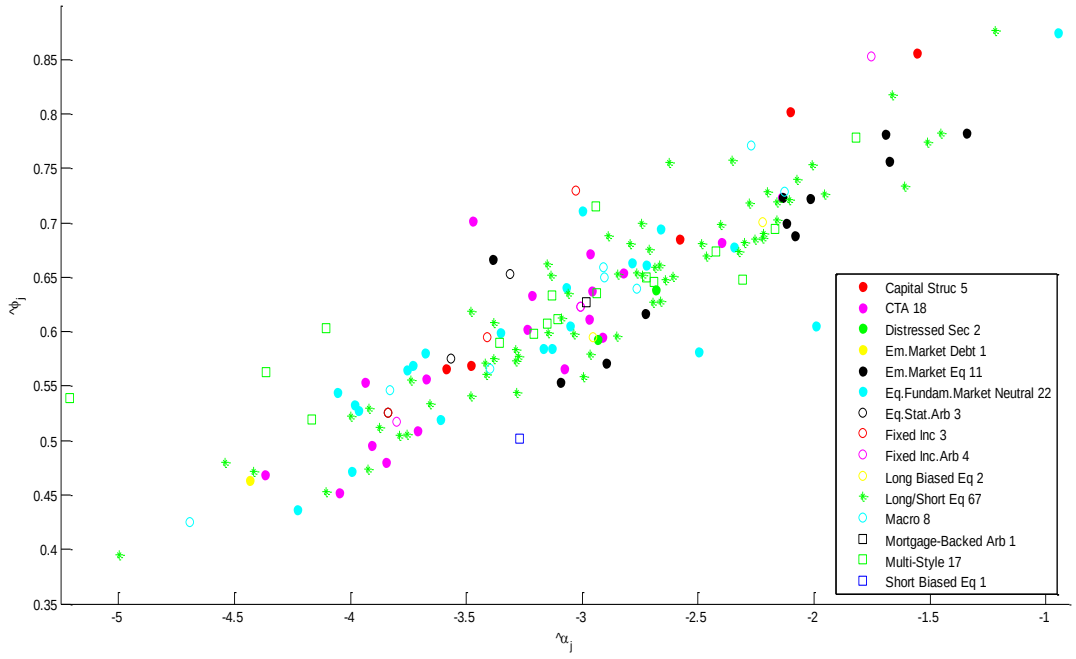


Figure 3.8: Estimated values of intercept (volatility) $\hat{\alpha}_{t,j}$ and persistence of volatility $\hat{\varphi}_{t,j}$ according to different management styles

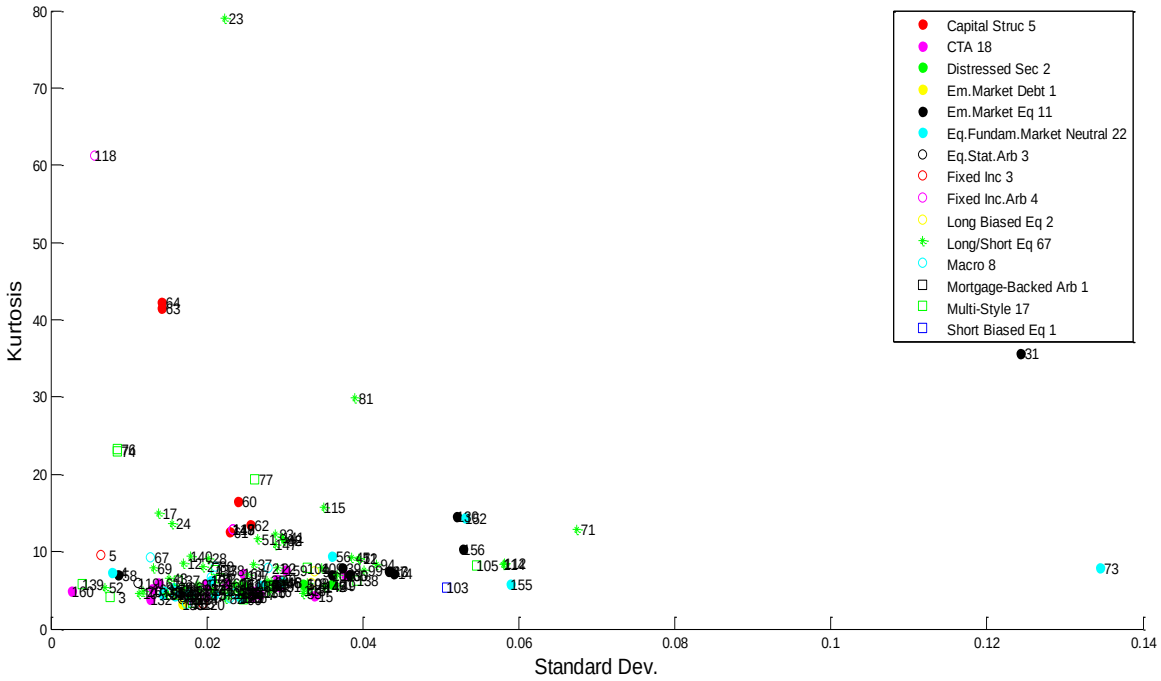


Figure 3.9 a): Standard deviation and kurtosis of individual hedge funds

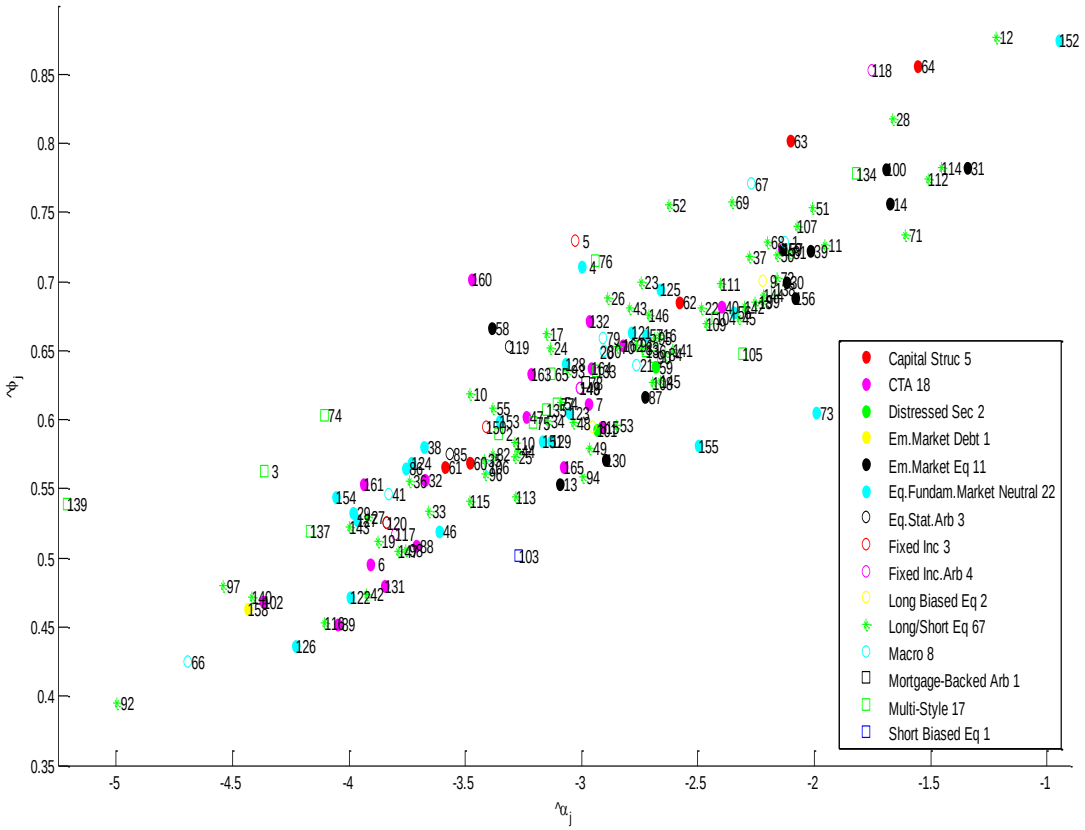


Figure 3.9 b): Estimated values of intercept (volatility) $\hat{\alpha}_{t,j}$ and persistence of volatility $\hat{\varphi}_{t,j}$.

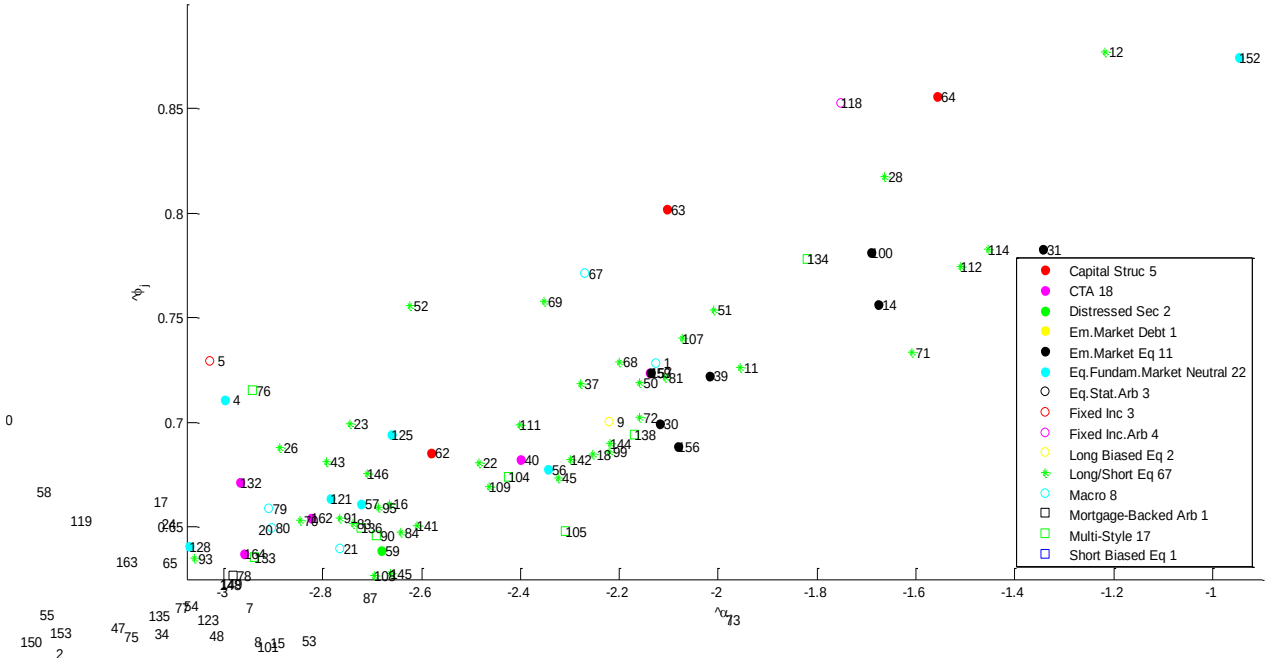


Figure 3.10: Zoom on the Figure 9b)

Chapter 4 - Conclusion

We propose a new Bayesian hierarchical SV model for joint analysis of hedge funds. We review the literature on the estimation of the SV models, emphasising the importance of the Bayesian paradigm. We discuss how simulation methods such as Gibbs sampling and Metropolis-Hastings algorithm apply to the SV models. We apply an inference method based on Bayesian approach and propose a MCMC estimation procedure. We reveal that MCMC method is effective in estimating the true value of the latent volatility process. Finally, we compare MCMC estimation of parameter $\hat{\varphi}_{t,j}$with kurtosis of individual hedge funds and we find that high kurtosis corresponds to the high level of persistence of volatility. We also find heterogeneity in the level of volatility and persistence across funds.

APPENDIX

The JB test statistic is defined as JB = $\frac{n-1}{6}\left(S^2 + \frac{(K-3)^2}{4}\right)$, where *n* is the sample size, *S* is the sample skewness and *K* is the sample kurtosis.

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^3}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{3/2}}$$

$$K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2\right)^{2}},$$

where $\hat{\mu}_3$ and $\hat{\mu}_4$ are the estimates of the third and fourth central moments, respectively, $\bar{x}$ is the sample mean, and $\hat{\sigma}^2$ is the estimate of the second central moment, the variance.

REFERENCES

Andersen, T.G., and Sorensen, B. (1996). GMM estimation of a stochastic volatility model: a Monte Carlo study. *Journal of Business and Economic Statistics,* 14, pp. 328 – 352.

Bansal, R., Gallant, A.R., Hussey, R., and Tauchen, G. (1993). Computational aspects of non parametric simulation estimation. In *Computational Techniques for Econometrics and Economic Analysis,* ed. by Belsley, D.A., Kluwer Academic Publishers, Boston, MA, pp. 3 – 22.

Bauwens, L., Lubrano, M., and Richard, J.-F. (1999). *Bayesian Inference in Dynamic Econometric Models*. 1st edition. Oxford University Press, Oxford.

Billio, M., Getmansky, M., and Pelizzon, L. (2009). Non-Parametric Analysis of Hedge Fund Returns: New Insights from High Frequency Data. *The Journal of Alternative Investment,* 12, pp. 21 – 38.

Billio, M., Getmansky, M., and Pelizzon, L. (2012). Dynamic Risk Exposure in Hedge Funds. *Computational Statistics and Data Analysis,* 56, pp. 3517 – 3532.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics,* 31, pp. 307–327.

Danielson, J. (1994). Stochastic volatility in asset prices: estimation with simulated maximum likelihood. *Journal of Econometrics,* 61, pp. 375 – 400.

Duffie, D., and Singleton, K.J. (1993). Simulated moments of Markov models of asset prices. *Econometrica,* 61, pp. 929 – 952.

Engle, R. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, *Econometrica*, 50, pp. 987-1008.

Fung, W., and Hsieh, D. (1997). Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds. *Review of Financial Studies,* 10, pp. 275 – 302.

Fung, W., and Hsieh, D. (2000). Performance Characteristics of Hedge Funds and CTA Funds: Natural versus Spurious Biases. *Journal of Financial and Quantitative Analysis,* 35, pp. 291-307.

Fung, W., and Hsieh, D. (2001). The risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers. *Review of Financial Studies,* 14, pp. 313 – 341.

Gallant, A.R., Hsieh, D., and Tauchen, G. (1997). Estimation of stochastic volatility models with diagnostics. *Journal of Econometrics,* 81, pp. 159 – 192.

Gelfand, A.E., and Smith, A.F.M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of Statistical Association,* 85, pp. 972-985.

Gelman, S., Carlin, J., Stern, H.S., and Rubin, D.B. (2003). *Bayesian Data Analysis,* 2nd edition. Chapman and Hall/CRC Press, London.

Geman, S., and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. PAMI,* 6, pp. 721-741.

Getmansky, M., Lo, A.W., and Makarov, I. (2004). An economic model of serial correlation and illiquidity in hedge fund returns. *Journal of Financial Economics,* 74, pp. 529 – 609.

Hastings, W.K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika,* 57, pp. 97-109.

Jacquier, E., Polson, N.G., and Rossi, P.E. (1994). *Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics,* 12, pp. 371–389.

Klaus, B., and Rzepkowski, B. (2009). Risk spillover among hedge funds. The role of redemptions and fund failures. *Working Paper Series 1112, European Central Bank.* Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1502467

KPMG International (April, 2012). The value of the hedge fund industry to investors, markets, and the broader economy. Available at https://www.kpmg.com/KY/en/Documents/the-value-of-the-hedge-fund-industry-part-1.pdf

Lopes, H., and Polson, N. (2010). Bayesian inference for stochastic volatility modelling, in *Rethinking Risk Measurement and Reporting: Uncertainty, Bayesian Analysis and Expert Judgement*, ed. by Bocker, K., pp. 515-551.

Metropolis, N., and Ulam S. (1949). The Monte Carlo Method, *Journal of American Statistical Association,* 44, pp. 335-341.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M. N., Teller, A.H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics,* 21, pp. 1087-1091.

Mitchell , M., and Pulvino, T. (2001) . Characteristics of Risk and Return in Risk Arbitrage. *Journal of Finance,* 56(6), pp. 2135-2175.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometric*a, 59, pp. 347-370.

Robert, C.P., and Casella, G. (1999). *Monte Carlo Statistical Methods.* Springer – Verlag New York.

Tanner, M.A., and Wong, W.H. (1987). The calculation of posterior distributions by data Augmentation (with discussion). *Journal of the American Statistical Association,* 82, pp. 528-550.

Tran, V. Q. (2006). *Evaluating Hedge Fund Performance.* 1st edition. John Wiley & Sons.

Tsay, R.S. (2005). *Analysis of Financial Time Series.* 2nd edition. John Wiley & Sons.

WEB SITE REFERENCES

HFR Industry reports:

https://www.hedgefundresearch.com/?fuse=products-irglo

HedgeCo.Net:

http://www.hedgeco.net/hedgeducation/industry-overview.php

Internet Journal dedicated to the Bayesian Inference:

http://www.bayesian-inference.com