ESSAYS ON BAYESIAN ANALYSIS OF FINANCIAL ECONOMICS

by

LIULING LI

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Economics

written under the direction of

Professor Hiroki Tsurumi

and approved by

____

____

____

____

New Brunswick, New Jersey

May, 2009

ABSTRACT OF THE DISSERTATION

Essays on Bayesian Analysis of Financial Economics

By LIULING LI

Dissertation Director:
Professor Hiroki Tsurumi

This dissertation consists of three essays with each essay forming a chapter. The regression models in these three chapters are different but share the same feature: the error terms of the models all follow ARMA-GARCH error processes generated either from normal or exponential power distributions.

In the first chapter I present a spot asset pricing model that is known as the CKLS model. Two CKLS models are compared. In one model the ARMA-GARCH error process is generated by the exponential power distribution while in the other model the error process is generated by the normal distribution. Using monthly U.S. federal funds rate I estimate the parameters of the CKLS models. From the predictive densities I obtain the distributions of the mean squared errors of forecast (MSEF) and the predictive deviance information criterion (PDIC). In addition I use the Bayes factor and the deviance information criterion (DIC). Markov Chain Monte Carlo (MCMC) algorithms, which are stochastic numerical integration methods, are used. I find that in general the CKLS model with the error term generated by the exponential power distribution is chosen over the model with the normal error term.

In the second chapter I first compare two MCMC algorithms: random walk draw and non-random walk draw for a Markov switching regression model. Two Markov switching models are compared: one with the variance of the normal distribution generated by the state space variable and the other with the constant variance. The realized volatilities of MMM Company are used to estimate and compare the models. The mean squared errors (MSE) and mean squared errors of forecast (MSEF) are used as the model selection criteria. I find that the model with the constant variance is chosen over the model with the state space variance by the MSE but the latter is chosen over the former by the MSEF.

In the third chapter I estimate a bivariate copula model. Each of the two regressions is generated by the exponential power distribution. I use monthly data on SP500 and FTSE100. Results show that the correlation parameter for SP500 and FTSE100 is .6893.

Acknowledgement

I thank all those people who have made this thesis possible and have made writing this dissertation such an enjoyable experience for me. I am deeply indebted to Professor Hiroki Tsurumi. He has always encouraged me to explore the topics I am interested in. I am grateful to Professor Norman Swanson for his useful suggestions. His lecture on time series econometrics inspired me deeply. I am greatly thankful to Professor Bruce Mizrach for his opinions. His lecture on financial economics expanded my knowledge horizon greatly. Thanks to Professor Elena Goldman for serving as the outside member in my dissertation defense and providing me data on the realized volatility. Acknowledgement is also made to Professor Roger Klein. His lecture on econometrics provided me a solid background for further study. Thanks are also due to Dorothy Rinaldi who is always available for answering any question during my Ph.D. study.

My thanks also go to the team of the economics department at Rutgers. I acknowledge the financial support and the mock interviews from the economics department. The team spirit at the economics department is excellent. My Ph.D. study is deeply supported by the team.

Last but not least, I would like to thank my family and friends. I am grateful to my parents, sisters, husband and son for their consistent support.

Table of Contents

List of tables

List of illustrations

1 Introduction

Better modeling and forecasting macroeconomic variables and financial time series such as interest rates, stock index returns and volatility are important issues in both finance and economics. In my dissertation, I focus on model designing, model comparison and model selection. My objective is to apply econometrics to financial economics. Bayesian inference is used to analyze time series in economics. The three chapters in my dissertation are as follows.

In the 1st chapter, I create a CKLS-ARMA-GARCH-EPD model for short-term interest rate. This model adds some new features to the CKLS model suggested by Chan, Karolyi, Longstaff and Sanders (1992). Firstly, the diffusion parameter follows a GARCH process to capture the high degree of volatility persistence (see Engle (1982) and Bollerslev (1987)). Secondly, exponential power distribution (EPD) (see Subbotin (1923)) is assumed for the error term to better model the fat tail property of interest rate.

Markov Chain Monte Carlo algorithm (MCMC) is used to estimate this model. Then, the predictability of this model is compared with other model. Bayes Factor (see Gelfand and Dey (1994)), Deviance Information Criterion (DIC) (see Spiegelhalter(2002)), Predictive Deviance Information Criterion (PDIC) and Mean Squared Error of Forecast are generated within the MCMC algorithms and used as model selection criteria.

The DIC shows CKLS-ARMA-GARCH-EPD has better in-sample fit. The long-run equilibrium value for short-term interest is 1.53% calculated from the CKLS-ARMA-GARCH-EPD. And out-of-sample forecast comparison shows the model with better in-sample fit is not necessary the one with better forecastibility. Rolling estimation scheme shows CKLS-ARMA-GARCH-EPD predicts long-run better.

One of the future extensions for chapter 1 is to add some macro variables of Bernake et al. (2005)) into CKLS-ARMA-GARCH-EPD model and then compare its forecastibility with other models. Another extension could add a regime switch parameter in CKLS-ARMA-GARCH-EPD model (see Dueker et al. (2007)). One can also add 2 or 3 latent factors in this model (see Sorwar(2005)). CKLS-ARMA-GARCH-EPD model can be used to price derivatives. Or we can perform some non-nested model comparison. Last but not the least, we can add jump components into this model.

In the 2nd chapter, we study Markov switching models by Bayesian inference. Markov switching model (hereafter denoted as MSW) was introduced to economics by Hamilton (1989) . We first compare two MCMC algorithms: random walk draw and non-random walk draw for a Markov switching model with state variable in the variance(i.e., MSW-ARMA-GARCH-NORMAL-S_t-in-variance, see Das and Yoo(2004)). Then, we compare this Markov switching model with other model (i.e., MSW-ARMA-GARCH-NORMAL-S_t-in-mean, see Yoo(2006)) by MSE and MSEF for the realized volatility of MMM stock prices. For realized volatility, one can refer to Andersen and Bollerslev (1997).

Results show algorithm of non-random walk draw is better. And MSW-

ARMA-GARCH-NORMAL-S_t-in-mean has better in-sample fit, but MSW-ARMA-GARCH-NORMAL-S_t-in-variance has better out-of-sample forecastibility. Future extensions for chapter 2 could be follows. First, we can compare Markov switching model with other nonlinear models such as TARMA. Secondly, we can consider other model selection criteria. Thirdly, we can also extend previous Markov switching models by assuming other error distributions such as the exponential power distribution. Last but not least, we can design a new model which includes two regime switching variables: one regime switching variable in the mean component, another regime switching variable in the variance component.

In chapter 3, we design a Gaussian copula model with ARMA-GARCH-EPD error terms (i.e., Copula-ARMA-GARCH-EPD). Copulas are convenient tools to construct multivariate joint distributions from margins. In financial risk assessment and actuarial analysis, dependence modeling with copula functions is widely used. Li (2000) introduces Gaussian copula model into pricing Collateralized Debt Obligations (CDOs). We design Markov Chain Monte Carlo (MCMC) algorithms to estimate the Copula-ARMA-GARCH-EPD model. Monthly data on SP500 and FTSE100 are analyzed.

Results show the MCMC algorithm is convergent. The correlation parameter ρ for SP500 and FTSE100 is .6893. One of the future extensions for chapter 3 is to compare Copula-ARMA-GARCH-EPD model with other nonlinear models. Another extension could add a regime switch parameter in Copula-ARMA-GARCH-EPD model (see Dueker et al. (2007)). Or we can add some macro variables of Bernake et al.(2005)) into Copula-ARMA-GARCH-EPD model and then compare its forecastibility with other models. Last but not the least, one may use Copula-ARMA-GARCH-EPD to price CDOs.

2 CKLS models for US short-term interest rate

2.1 Motivation

A short term interest rate is one of the most important economic indicators. Better modeling and forecasting the short term interest rate are crucial for risk management, derivatives pricing and monetary policy.

Since the mid 1970's, asset prices have been modeled by diffusion processes. In 1977, Vasicek(1977) applied an Ornstein–Uhlenbeck process to model the short term interest rate. Since the Vasicek model can generate negative interest rates, Cox, Ingersoll and Ross(1985) suggested a square root process. Chan, Karolyi, Longstaff and Sanders(1992) proposed a more flexible model now known as the CKLS model. The CKLS model nests a class of asset pricing models as shown in Table 1.

In CKLS formula, r_t is the asset price at time t or interest rate level. W_t

CKLS(1992)	$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^{\gamma} dW_t$			
Merton(1973)	$dr_t = \alpha dt + \sigma dW_t$		$\beta = 0$	$\gamma = 0$
Black and Scholes(1973)	$dr_t = \beta r_t dt + \sigma r_t dW_t$	$\alpha = 0$		$\gamma = 1$
Cox (1975)	$dr_t = \beta r_t dt + \sigma r_t^{\gamma} dW_t$	$\alpha = 0$		
Vasicek(1977)	$dr_t = (\alpha + \beta r_t)dt + \sigma dW_t$			$\gamma = 0$
Dothan(1978)	$dr_t = \sigma r_t dW_t$	$\alpha = 0$	$\beta = 0$	$\gamma = 1$
Brennan and Schwartz(1980)	$dr_t = (\alpha + \beta r_t)dt + \sigma r_t dW_t$			$\gamma = 1$
Cox et al.(1980)	$dr_t = \sigma r_t^{1.5} dW_t$	$\alpha = 0$	$\beta = 0$	$\gamma = 1.5$
Cox et al.(1985)	$dr_t = (\alpha + \beta r_t)dt + \sigma r_t^{1/2} dW_t$			$\gamma = .5$

Table 1: Asset Pricing Models Nested in the CKLS Model

is a Wiener Process (or Brownian motion), α, β, σ and γ are parameters. β measures the speed of mean reversion. r_t mean-reverts towards the level α/β. σ is the diffusion parameter. γ is the constant elasticity of variance with respect to interest rate level.

The CKLS model has been used in many studies. For example, Sanford(2006) studied Australia interest rate using CKLS. And Antoniou et al.(2006) compared different models for UK interest rate. Other studies that use the CKLS model are Tse(1995), Dahlquist(1996), Gray(1996), Nowman(1997), Jiang and Knight(1997), Hiraki and Takezwa(1997), Byers and Nowman(1998), Brailsford and Maheswaran(1998), Ball and Torous(1999), Dennis et al.(1999), Saltoğlu(2003), Nowman(2003), Treepongkaruna and Gray(2003), Rodrigues(2005), Goldman(2005), Sorwar(2007), Qian et. al.(2008) and Nowman(2008).

There are two limitations of the CKLS model. One is the diffusion parameter[1] σ of the CKLS model is a constant, which may not capture the extremely high degrees of volatility in the asset prices. The other limitation is the CKLS model discretized by Euler-Maruyama scheme assumes normal distribution for the error term. The discrete time specification of the CKLS

[1]For discrete-time specification of the CKLS model, σ is also the parameter in variance if variance exists or is finite.

model by Euler-Maruyama scheme[2] is

$$r_t - r_{t-1} = \alpha + \beta r_t + \sigma r_t^{\gamma} \varepsilon_t \ , \varepsilon_t \sim N(0,1) \tag{1}$$

To overcome the limitations of CKLS model and better model the high degrees of volatility persistence, we estimate the CKLS model with an ARMA-GARCH error term that follows an exponential power distribution (denoted as CKLS-ARMA-GARCH-EPD or Model 1). The reasons we use GARCH for conditional variance are 1) in the literature of finance and economics, the ARCH model of Engle(1982) and the GARCH model of Bollerslev(1986) have been used widely to model the high degrees of volatility persistence. 2) Brenner et al.(1996) and Koedijk et al.(1997) show the CKLS-GARCH model performs much better statistically than either the CKLS or the GARCH models with white noise errors.

Even though normal distribution assumption has been widely used in many models, there are non-normal distributions used in literature. For example, Mandelbrot(1963) and Fama(1963) used the stable Paretian family of distribution. Bollerslev(1987) extended the GARCH model to allow for conditionally t-distributed errors. Liu and Brorsen(1995) assumed a stable distribution for the error term of a GARCH model with application to foreign currency returns. Nelson(1991) and Bali and Wu(2006) used the generalized

[2] To discretize a continuous-time model, we can use Euler-Maruyama scheme, Milstein scheme or other more sophisticated approximation techniques. We choose Euler-Maruyama scheme for simplicity.

error distribution(or exponential power distribution, i.e., EPD[3]) for the error term[4]. We employ the exponential power distribution (EPD) for the error in our model. In addition, we specify that the error term follows an ARMA-GARCH model.

There are three ways to estimate the diffusion models: GMM, MLE, and Bayesian. Aït-Sahalia(1996) estimated the diffusion interest rate model using GMM. Conley, Hansen, Luttmer, and Scheinkman(1997) estimated the parameter γ by 2-step GMM. For more information about GMM or other nonparametric estimations, please see Chan et al.(1992), Longstaff and Schwartz(1992), Aït-Sahalia(1996), Bates(1996), Conley et al.(1997), Andersen and Lund(1997), Stanton(1997), Jiang(1998), Ahn(1999), Ahn and Gao(1999), Singleton(2001), Cont and Tankov(2002), Bandi and Nguyen(2003), Chernov et al.(2003), Fan(2003), Bhardwaj(2005), Renò(2005), and Faff and Gray(2006). One disadvantage of the GMM is it may not allow the estimation of the unobserved volatility dynamics. Also, researches show some problems with GMM. Eom(1998) found that the choice of moment conditions in GMM have great impacts on the parameter estimates. Broze, Scaillet and Zakoian (1995) showed when $\gamma > 1$ in the CKLS model, the GMM estimator is not well behaved. Faff and Gray(2006) found out there are severely biased

[3] EPD is also known as the generalized error distribution. The EPD is proposed by Subbotin (1923). For more information about EPD, please see Box and Tiao (1973), and Tsurumi and Shimizu (2008).

[4] For more information about other error distribution assumptions, see Praetz(1972), Blattberg and Gonedes(1974), Press(1967), Clark(1973), Kon(1984), Fielitz and Rozelle(1983) and Boothe and Glassman(1985).

in the GMM estimates of the speed of mean reversion β in CKLS model. Qian et al.(2005) showed the Bayesian and MLE perform better than GMM.

The second is the maximum likelihood method (MLE) including quasi MLE. Aït-Sahalia(2002) used a closed-form approximation for the likelihood function and then used maximum likelihood method to estimate the discretely sampled diffusion. For more information about MLE, see Marsh and Rosenfeld(1982), Nelson(1991), Brenner et al. (1996), Koedijk et al.(1997), Aït-Sahalia(1999, 2002), Bali(2000), Das(2002), Jagannathan et al.(2003), Demirtas(2006), Bali and Wu(2006), Dueker et al.(2007), and Christiansen(2008). The disadvantage of MLE is this method is restrictive because few short term interest rate models have known or simple likelihood functions.

The third is the Bayesian method, i.e. Markov Chain Monte Carlo method (MCMC), see Elerian et al.(2001), Eraker(2001, 2004), Jones(2003), Hong(2005), Goldman(2005), Sorwar (2005), Griffin and Steel (2006), Sanford et al.(2006), Li et al.(2006) and Qian et al.(2005). The advantages of MCMC are that it can estimate the posterior distributions of the parameters easily; it can estimate the unobserved volatility; it can be used to estimate models with complicate posterior distributions and it can also estimate the predictive density for the forecast value.

In this paper, we use Bayesian method (MCMC). The in-sample fit of this new model is analyzed by comparing with CKLS-ARMA-GARCH-NORMAL. Bayes Factor (see Gelfand and Dey (1994)) and Deviance Information Criterion (DIC, see Spiegelhalter(2002)) are used as in-sample fit criteria. The long run equilibrium of short-term interest rate is calculated.

Lastly, we also test one hypothesis: if one model has better in-sample fit, will it also be the model with better forecastibility? Predictive Deviance Information Criterion (PDIC) and the distributions of Mean Squared Error of Forecast (MSEF) are generated within the Markov Chain Monte Carlo algorithms and used as out-of-sample model selection criteria.

Simulation results show that the MCMC algorithms used to estimate and forecast CKLS-ARMA-GARCH-EPD and CKLS-ARMA-GARCH-NORMAL are efficient since the filtered Fluctuation test (FT) or the filtered Kolmogorov Smirnov test (KS) (see Ploberger, Kramer and Kontrus(1989), Smirnov(1939) and Goldman et.al. (2007)) indicates convergence.

The results of in-sample model comparison are as follows. DIC shows CKLS-ARMA-GARCH-EPD model fits in-sample data better since its DIC value is smaller. However, Bayes Factor has decisive evidence supporting Model 2. DIC uses effective number of parameters instead of the true number of parameters to compare models. So, we prefer to use the DIC criterion and pick CKLS-ARMA-GARCH-EPD model as the one with better in-sample fit. And the long-run equilibrium short-term interest rate is 1.53%, which is calculated using the posterior means of parameter estimates from CKLS-ARMA-GARCH-EPD model.

The results of forecast show a model with better in-sample fit does not necessary have better out-of-sample forecastibility. One-period ahead PDIC shows CKLS-ARMA-GARCH-NORMAL has better out-of-sample forecast since its PDIC value is smaller. If we use one-period ahead Mean Squared Error of Forecast(MSEF), CKLS-ARMA-GARCH-EPD is better. Since MSEF criterion comes from sample theory and PDIC criterion comes from Bayesian theory, we prefer the results of PDIC.

To test the robustness of previous results, 3-period, 6-period and 9-period ahead forecast values (fixed estimation scheme) are also calculated. The results of one-period ahead PDIC and MSEF are robust. For PDIC, CKLS-ARMA-GARCH-NORMAL model always has better forecastibility since its PDICs are smaller. For MSEF, CKLS-ARMA-GARCH-EPD model has better forecastibility since its MSEFs are smaller. When out-of-sample period increases from 1 to 9, the MSEF of CKLS-ARMA-GARCH-NORMAL increases greatly.

To check the robustness of the results for PDIC from previous fixed estimation scheme, we use rolling estimation scheme to derive PDIC for 10-period, 15-period, 20-period and half-sample ahead data. The results of rolling estimation scheme show CKLS-ARMA-GARCH-EPD model has better forecastibility. The PDIC for CKLS-ARMA-GARCH-EPD model becomes smaller as the forecast period increases from 10 to half-sample.

This paper is organized as follows. Section 2 introduces the models and methodology. The effectiveness of the MCMC algorithms is shown in section 3. Section 4 shows the empirical results for model estimation and predictability. Conclusions are given in section 5.

2.2 CKLS-ARMA-GARCH models and MCMC algorithms

We specify the CKLS-ARMA-GARCH model as:

$$r_t = a + br_{t-1} + r_{t-1}^c u_t \tag{2}$$

where

$$
\begin{aligned}
u_t &= \sum_{j=1}^{p} \phi_j u_{t-j} + \sum_{j=1}^{q} \theta_j e_{t-j} + e_t && (3)\\
e_t &= \sigma_t \varepsilon_t && (4)\\
\sigma_t^2 &= \alpha_0 + \sum_{i=1}^{r} \alpha_i e_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}^2 && (5)\\
\alpha_0 &> 0, \alpha_i \geq 0,\ i = 1, ..., r.\ \beta_i \geq 0\ , i = 1, ..., s. && (6)\\
1 &\geq \sum_{i=1}^{\max(r,s)} (\alpha_i + \beta_i) && (7)
\end{aligned}
$$

r_t is spot rate. The error term ε_t follows the exponential power distribution (EPD) with variance normalized to be unity. The probability density function of ε_t is given by

$$
\begin{aligned}
f(\varepsilon_t|Y, X) &= \frac{1}{\lambda 2^{1+1/\alpha}\ \Gamma(1 + 1/\alpha)} \exp\left\{-\frac{1}{2}|\frac{\varepsilon_t}{\lambda}|^{\alpha}\right\} && (8)\\
\lambda &= \sqrt{\frac{2^{-2/\alpha}\Gamma(1/\alpha)}{\Gamma(3/\alpha)}} && (9)
\end{aligned}
$$

λ is the constant to make the variance of ε_t unity. The EPD distribution has been used in many studies (see Box and Tiao(1973), Nelson(1991) and Tsurumi and Shimizu(2008), among others). Let us call equation 2 with the EPD error term as Model 1 or CKLS-ARMA-GARCH-EPD model.

Majority of the CKLS models in the literature employ the normal error term: $\varepsilon_t \sim N(0,1)$. Let us label the model with the normal error term Model 2 or CKLS-ARMA-GARCH-NORMAL model. We shall compare the performance of Model 1 and Model 2 to see which model fits in-sample data better and which one performs better judged by forecast performance.

2.2.1 Densities of CKLS-ARMA-GARCH models

We use the posterior density as our target density in the MCMC algorithms to draw the parameter estimates. After we estimate these models, we compare these models using in-sample selection criteria such as Bayes Factor and Deviance Information Criterion. Also, we calculate the forecast values using the posterior means of parameters. Out-of-sample comparisons are made by Predictive Deviance Information Criterion (PDIC) and Mean Squared Error of Forecast (MSEF).

Assume $\Theta = \{a, b, c, \{\phi_i\}_{i=1\ldots p}, \{\theta_i\}_{i=1\ldots q}, \{\alpha_i\}_{i=0\ldots r}, \{\beta_i\}_{i=1\ldots s}, \alpha\}$ for Model 1, the posterior distribution for Y is:

$$p(\Theta|Y,X) \propto \prod_{t=1}^{T} \frac{r_{t-1}^{-c}\sigma_t^{-1}}{\lambda 2^{1+1/\alpha}\,\Gamma(1+1/\alpha)} \exp\left\{-\frac{1}{2}|\frac{\varepsilon_t}{\lambda}|^{\alpha}\right\}\ p(\Theta) \qquad (10)$$

where

$$\varepsilon_t = \frac{e_t}{\sigma_t} \tag{11}$$

$$e_t = \frac{r_t - (a + br_{t-1})}{r_{t-1}^c} - \sum_{j=1}^{p} \phi_j u_{t-j} - \sum_{j=1}^{q} \theta_j e_{t-j} \tag{12}$$

$$\lambda = \sqrt{\frac{2^{(-2/\alpha)}\Gamma(1/\alpha)}{\Gamma(3/\alpha)}} \tag{13}$$

And

$$\begin{aligned} p(\Theta) &= p\{a, b, c, \{\phi_i\}_{i=1\ldots p}, \{\theta_i\}_{i=1\ldots q}, \{\alpha_i\}_{i=0\ldots r}, \{\beta_i\}_{i=1\ldots s}, \alpha\} \\ &= p(a)p(b)p(c)\prod_{i=1}^{p} p(\phi_i)\prod_{i=1}^{q} p(\theta_i)\prod_{i=0}^{r} p(\alpha_i)\prod_{i=1}^{s} p(\beta_i)p(\alpha) \\ &= N_a(0,\ .1)N_b(0,\ .1)Nc(0,\ .1)\prod_{i=1}^{p} N_{\phi_i}(0,\ .1)\prod_{i=1}^{q} N_{\theta_i}(0,\ .1) \\ &\quad \prod_{i=0}^{r} N_{\alpha_i}(0,\ .1)\prod_{i=1}^{s} N_{\beta_i}(0,\ .1)N_\alpha(0,\ .1) \end{aligned} \tag{14}$$

is the prior. We assume priors for each parameters are independent. For model 2, we set parameter $\alpha = 2$.

2.2.2 MCMC algorithms

In order to implement this model, for parameter (α_i, β_i), we follow the approximation of Nakatsuma(1998) to get proposal density:

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^{l}(\alpha_i + \beta_i)\varepsilon_{t-i}^2 - \sum_{i=1}^{s} \beta_i w_{t-i} + w_t,\ w_t \sim N(0, 2\sigma_t^2) \tag{15}$$

Where $w_t = \varepsilon_t^2 - \sigma_t^2$, $l = max(r, s)$, $\alpha_i = 0$ for $i > r$, $\beta_i = 0$ for $i > s$, $\varepsilon_t^2 = 0$ and $w_t = 0$ for $t \leq 0$.

Nakatsuma(1998) used nonlinear least square estimation to draw parameters in the MA and GARCH processes. Qian et al.(2005) used random walk draws for all parameters. We follows Qian et al.(2005)'s modifications.

Different from Nakatsuma(1998), we need to draw both parameter c in the CKLS model and the parameter α in EPD. Different from Qian et al.(2005), we use Efficient Jump Algorithms to draw parameter $\{\theta_i\}_{i=1...q}$ and $\{\beta_i\}_{i=1...s}$, and we need to draw parameter α in EPD.

Nakatsuma(1998) and Qian et al.(2005) focus on parameter estimation. However, we not only estimate parameters in Θ but also calculate the predictive values $\widetilde{Y}$ using the posterior means of parameters. In addition, we focus on model selection.

In our MCMC algorithm, we also compare following two algorithms for drawing parameter α:

1) Efficient Jump Algorithm.

2) Modified Efficient Jump Algorithm with Inverse Gaussian as the proposal density.

2.3 Simulation results

Simulation data we used are:

$$
\begin{aligned}
a &= .2,\ b = .4,\ c = .5 \\
\phi_1 &= .4,\ \theta_1 = .4 \\
\alpha_0 &= .4,\ \alpha_1 = .7,\ \beta_1 = .1 \\
\alpha &= 1
\end{aligned}
$$

The sample size is T=1000. We run 18000 MCMC iterations. And we discard the first 3000 draws and keep every 10th draws. All acceptance rates are higher than 0.28.

We check the convergence of the draws by plotting the draws, by the filtered Fluctuation test or filtered Kolmogorov-Smirnov test (see Ploberger, Kramer and Kontrus(1989), Smirnov(1939) and Goldman et.al.(2007)). We judge convergence by 3 criteria: 1) eye rolling convergence. i.e., the draws plotted in graphs are convergent. 2) P-value of filtered Fluctuation test is no less than 5% significance level. 3) or P-value of filtered Kolmogorov-Smirnov test is no less than 5% significance level. If any one of these 3 criteria is satisfied, we conclude the draws are convergent.

Simulation results show the draws from our MCMC algorithms are convergent. The plots of the draws indicate convergence[5]. And the convergence tests also indicate convergence(see Table 2). In Table 2, for each parameter,

[5] Graphs will be availabe upon request.

either the P-value of filtered Fluctuation test or filtered Kolmogorov-Smirnov test is no less than 5% significance level.

The summary statistics of the MCMC draws are given in Table 3. We can see the posterior means of parameters drawn from our MCMC algorithm are very close to the true values.

The α draws from model 1 using simulated data are plotted in Figure 1, 2, 3 and 4. We can see the draws of α are convergent and the mean of α is around 1, which is the true simulation parameter for Laplace distribution. The results from Efficient Jump Algorithm (see Figure 3 and 4) are similar to those from the Modified Efficient Jump with Inverse Gaussian as proposal density (see Figure 1 and 2).

2.4 Application to US federal funds rate

In the literature, there are many proxies for short-term interest rates. Some use monthly data (see Longstaff and Schwartz(1992), Chan et al.(1992), Ahn(1999), Nowman(1997), Aït-Sahalia(1999), Ahn and Gao(1999), and Brenner et al.(1996)). Others use weekly data (see Andersen and Lund(1997), Bali(2000), Eraker(2001), Jagannathan et al.(2003), Jones(2003), Andersen et al. (2005), Bhardwaj(2005), Demirtas(2006), Bali and Wu(2006), and Christiansen(2008)). Daily data are used in Conley et al.(1997), Das(2002),

Parameters	P-value of FT	P-value of KST
Model 1: CKLS-ARMA-GARCH-EPD		
a	0	.84
b	.52	.26
ϕ_1	.46	0
θ_1	.35	.03
α_0	.27	0
α_1	.19	.72
β_1	.12	.93
c	.31	.17
α	.05	.98
Model 2: CKLS-ARMA-GARCH-NORMAL		
a	0	.87
b	.44	0
ϕ_1	.43	0
θ_1	.34	.95
α_0	.11	.67
α_1	.26	.42
β_1	.12	.84
c	.19	.54

Table 2: P-values of Filtered Fluctuation Tests and Filtered Komogorov-Smirnov Tests

	True Value	Mean	St.Dev.
Model 1: CKLS-ARMA-GARCH-EPD			
a	.2	.1984	.03
b	.4	.4182	.11
ϕ_1	.4	.2182	.14
θ_1	.4	.4889	.26
α_0	.4	.5042	.09
α_1	.7	.8499	.08
β_1	.1	.0768	.07
c	.5	.4020	.18
α	1	.8843	.19
Model 2 : CKLS-ARMA-GARCH-NORMAL			
a	.2	.2288	.05
b	.4	.5360	.10
ϕ_1	.4	.2289	.06
θ_1	.4	.5065	.25
α_0	.4	.4051	.06
α_1	.7	.6320	.10
β_1	.1	.1934	.12
c	.5	.4448	.16

Table 3: Posterior Summaries

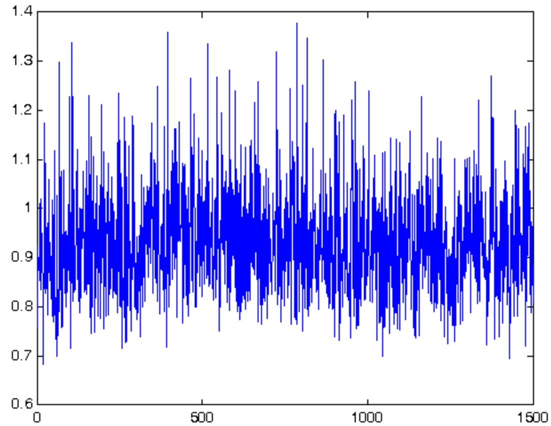

Figure 1: MCMC Draws of Parameter α Using Algorithm of Modified Efficient Jump with Inverse Gaussian as Proposal Density

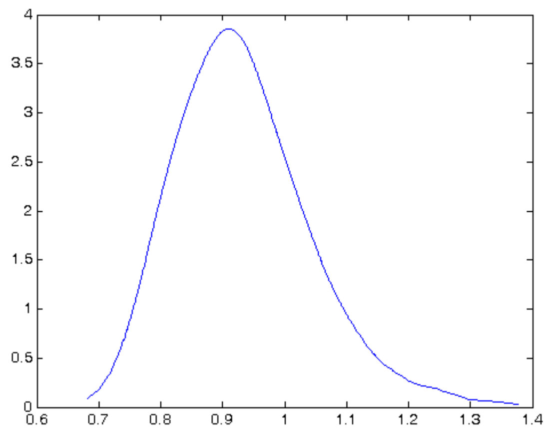

Figure 2: PDF of MCMC Draws of α Using Algorithm of Modified Efficient Jump with Inverse Gaussian as Proposal Density

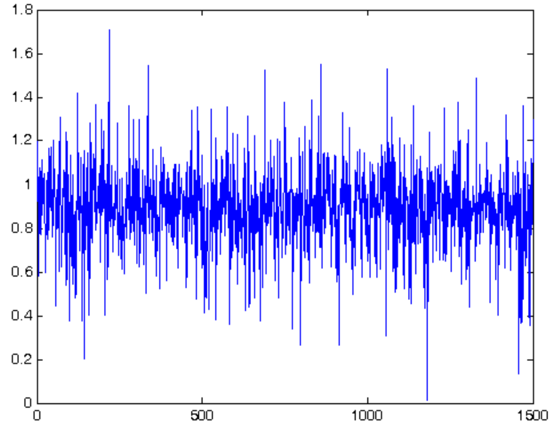

Figure 3: MCMC Draw of Parameter α Using Algorithm of Effecient Jump

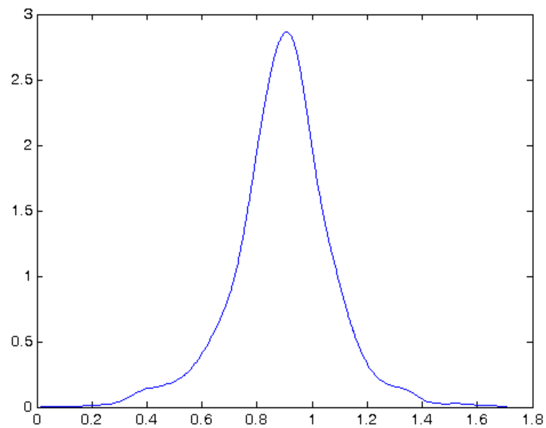

Figure 4: PDF of MCMC Draws of α Using Algorithm of Efficient Jump

Eraker et al.(2003), Fan(2003), and Qian et al.(2008). Quarterly data are used in Dueker et al.(2007).

2.4.1 Descriptive statistics

Monthly US Effective Federal Funds Rate is used to estimate the CKLS-ARMA-GARCH-EPD model[6]. And the results are compared with those from the CKLS-ARMA-GARCH-NORMAL model. The data are got from Board of Governors of the Federal Reserve System. The in-sample data consist of monthly Federal funds rate between Jan. 1963 – Dec. 2007. The out-of-sample data consist of monthly Federal funds rate between Jan. 2008 – Oct. 2008.

The descriptive statistics are listed in Table 4. The kurtosis is 5.2, which is greater than 3. And the skewness is 1.22, which is skewed to the right.

We also plot the unconditional density of the Federal funds rate in Figure 5. Compared to standard normal density, the unconditional density of the Federal funds rate is leptokurtic.

2.4.2 Parameter estimates and long-run equilibrium interest rate

The posterior means and standard deviations of the parameters are given in Table 5 for all models. The density for α is plotted in Figure 6. We can see the posterior mean of α is approximately .8424, far from Normal case $\alpha = 2$.

[6]Chan et al.(1992) use monthly US Treasury bill yield between June, 1964 and December, 1989 (306 observations) in their paper to estimate the CKLS model by GMM.

	R
Mean	6.30
Min	0.98
Max	19.10
Std.Dev.	3.26
Skewness	1.22
Kurtosis	5.20

Table 4: Descriptive Statistics for Federal Funds Rate

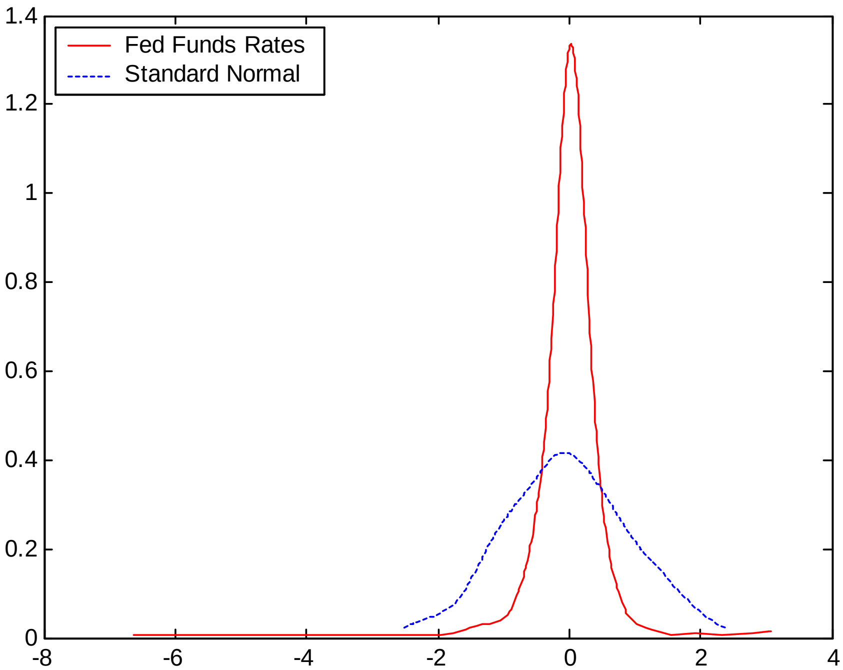

Figure 5: Unconditional Density of Federal Funds Rate

	Mean	St. Dev.
Model 1: CKLS-ARMA-GARCH-EPD		
a	2.8913	.15
b	$-.8882$	1.15
ϕ_1	.9980	.01
θ_1	.4593	.28
α_0	.0066	0
α_1	.7159	.15
β_1	.1479	.11
c	.6788	.34
α	.8424	.23
Model 2: CKLS-ARMA-GARCH-NORMAL		
a	2.9169	.13
b	$-.5145$	1.15
ϕ_1	.9888	.01
θ_1	.4994	.26
α_0	.0060	0
α_1	.5572	.18
β_1	.2194	.15
c	.6178	.27

Table 5: Posterior Means and Standard Deviations of the Parameters

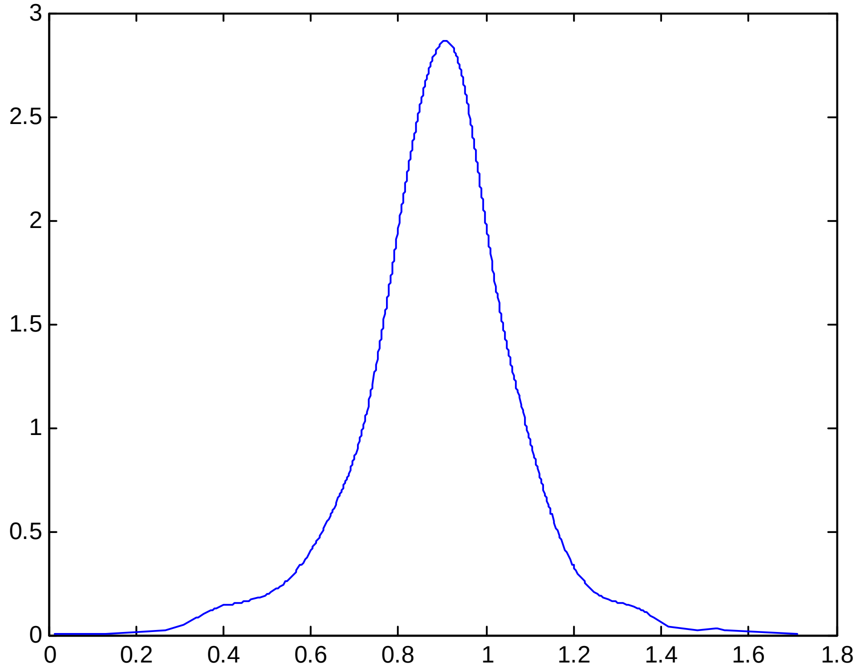

Figure 6: PDF of α Draws Using Interest Rate Data and Efficient Jump Algorithm

And the long-run equilibrium short-term interest rate is 1.53%, which is calculated using the posterior means of parameters a and b by following formula:

$$\frac{a}{-(b-1)} = \frac{2.8913}{-(-.8882-1)} = 1.53$$

In the paper of CKLS(1992), the long-run equilibrium short-term interest rate is 6.89%, which is calculated using the data in Table III of their paper (page 1218 of CKLS(1992)). We think the difference is caused by different models, different estimation techniques, and different data sets. Chan et al.(1992) use monthly US Treasury bill yield between June, 1964 and December, 1989 (306 observations) in their paper to estimate the CKLS model by GMM. But we use monthly US effective federal funds rate to estimate the CKLS-ARMA-GARCH-EPD model by MCMC.

2.4.3 In-sample-fit model selection criteria and results

To compare CKLS-ARMA-GARCH-EPD with CKLS-ARMA-GARCH-NORMAL, we first want to see which model has better in-sample fit. The model selection criteria are the Bayes Factor (see Gelfand and Dey (1994)) and the Deviance Information Criterion (see Spiegelhalter(2002)).

For Bayes Factor, the formula can be simplified as follows (see Gelfand

and Dey (1994), equation 21 on page 510):

$$\begin{aligned}\log PSBF &\approx \log \lambda_n + \frac{p_2 - p_1}{2} \\ &= \log \frac{L(\widehat{\theta}_1; Y, M_1)\pi_1(\widehat{\theta}_1)}{L(\widehat{\theta}_2; Y, M_2)\pi_2(\widehat{\theta}_2)} + \frac{p_2 - p_1}{2}\end{aligned} \tag{16}$$

We get Bayes Factor by following implementation:

$$\log \widehat{PS}BF = \log \frac{\frac{1}{nrep}\sum_{i=1}^{nrep} \widehat{L}(\widehat{\theta}_1^{(i)}; Y, M_1)\pi_1(\widehat{\theta}_1^{(i)})}{\frac{1}{nrep}\sum_{i=1}^{nrep} \widehat{L}(\widehat{\theta}_2^{(i)}; Y, M_2)\pi_2(\widehat{\theta}_2^{(i)})} + \frac{p_2 - p_1}{2} \tag{17}$$

where $\widehat{L}(\widehat{\theta}_1^{(i)}; Y, M_1)$ is the estimated likelihood function at *ith* draw for all parameter $\theta_1^{(i)}$ in Model 1. $\widehat{L}(\widehat{\theta}_2^{(i)}; Y, M_2)$ is the estimated likelihood function at the *ith* draw for all parameter $\theta_2^{(i)}$ in Model 2. $\pi_1(.)$ and $\pi_2(.)$ are priors. p_i is the number of parameters to be estimated in model i. $nrep$ is the number of accepted draws.

The Deviance Information Criterion (DIC) is suggested by Spiegelhalter(2002). The formula of DIC is as follows (see Spiegelhalter(2002), page 603, equation 36, 37):

$$DIC \equiv D(\overline{\theta}) + 2P_D = \overline{D(\theta)} - P_D \tag{18}$$

We calculate DIC by following implementation [7]:

$$
\begin{aligned}
\widehat{DIC} &\approx \widehat{\overline{D(\theta)}} - \widehat{P_D} \\
\widehat{\overline{D(\theta)}} &\approx -2 * \frac{1}{nrep}\sum_{i=1}^{nrep} \log L(y|\theta^{(i)}) \\
\widehat{D(\overline{\theta})} &\approx -2 * \log L(y|\overline{\theta}) \\
\widehat{P_D} &\approx \widehat{\overline{D(\theta)}} - \widehat{D(\overline{\theta})}
\end{aligned} \tag{19}
$$

Where $\overline{\theta}$ is the posterior mean of parameters. $Y = \{r_t\}_{t=1}^{T}$, and X are in-sample data. P_D is the effective number of parameters. $D(\theta)$ is the deviance information. $L(.)$ is the likelihood function:

$$
L(\Theta|Y, X) = f(Y|\Theta, X) \tag{20}
$$

$$
\propto \prod_{t=1}^{T} \frac{r_{t-1}^{-c}\sigma_t^{-1}}{\lambda 2^{1+1/\alpha}\,\Gamma(1+1/\alpha)} \exp\left\{-\frac{1}{2}|\frac{\varepsilon_t}{\lambda}|^{\alpha}\right\} \tag{21}
$$

$$
\begin{aligned}
\varepsilon_t &= \frac{e_t}{\sigma_t} \\
e_t &= \frac{r_t - (a + b r_{t-1})}{r_{t-1}^{c}} - \sum_{j=1}^{p}\phi_j u_{t-j} - \sum_{j=1}^{q}\theta_j e_{t-j} \\
\lambda &= \sqrt{\frac{2^{(-2/\alpha)}\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}
\end{aligned}
$$

The values of in-sample model selection criteria are listed in Table 6. DIC criterion shows CKLS-ARMA-GARCH-EPD (i.e., Model 1) fits in-sample data better since its DIC value is smaller. But Bayes Factor shows Model 2 is decisively better than Model 1. DIC uses effective number of parameters

[7]For more information about DIC implementation, one can also refer to Chen(2008).

P_D instead of the true number of parameters p_i to compare models. So, we prefer to use DIC criterion and pick CKLS-ARMA-GARCH-EPD model as the model with better in-sample fit.

2.4.4 Out-of-sample forecast model selection criteria and results

To see which model has better out-of-sample forecastibility, we use the Predictive Deviance Information Criterion (PDIC) and Mean Squared Errors of Forecast (MSEF).

We calculate the Predictive Deviance Information Criterion (PDIC) by the same formula of DIC using forecast values $\widehat{y}$. The formula of PDIC is as follows

$$\begin{aligned}
\widehat{DIC} &\approx \widehat{\overline{D(\theta)}} - \widehat{P_D} \\
\widehat{\overline{D(\theta)}} &\approx -2 * \frac{1}{nrep} \sum_{i=1}^{nrep} \log L(\widehat{y}|\theta^{(i)}) \\
\widehat{D(\overline{\theta})} &\approx -2 * \log L(\widehat{y}|\overline{\theta}) \\
\widehat{P_D} &\approx \widehat{\overline{D(\theta)}} - \widehat{D(\overline{\theta})}
\end{aligned} \tag{22}$$

Where $\overline{\theta}$ is the posterior mean of parameters. $\widehat{\mathbf{y}} = \{\widehat{r}_t\}_{t=1}^{M}$ is the forecast values. $\widetilde{\mathbf{x}}$ is out-of-sample data. $\widehat{\sigma}$ is the forecast values for conditional standard deviation. P_D is the effective number of parameters. $D(\theta)$ is the deviance information. $L(.)$ is the likelihood function:

$$
\begin{aligned}
L(\Theta|\widehat{Y},\widetilde{X}) &= f(\widehat{Y}|\Theta,\widetilde{X}) && (23)\\
&\propto \prod_{t=1}^{M}\frac{\widehat{r}_{t-1}^{-c}\widehat{\sigma}_t^{-1}}{\lambda 2^{1+1/\alpha}\,\Gamma(1+1/\alpha)}\exp\left\{-\frac{1}{2}|\frac{\widehat{\varepsilon}_t}{\lambda}|^{\alpha}\right\} && (24)\\
\widehat{\varepsilon}_t &= \frac{\widehat{e}_t}{\widehat{\sigma}_t}\\
\widehat{e}_t &= \frac{\widehat{r}_t-(a+b\widehat{r}_{t-1})}{\widehat{r}_{t-1}^{c}}-\sum_{j=1}^{p}\phi_j\widehat{u}_{t-j}-\sum_{j=1}^{q}\theta_j\widehat{e}_{t-j}\\
\lambda &= \sqrt{\frac{2^{(-2/\alpha)}\Gamma(1/\alpha)}{\Gamma(3/\alpha)}}
\end{aligned}
$$

We also calculate Mean Squared Errors of Forecast (MSEF):

$$
MSEF=\frac{1}{m}\sum_{i=1}^{m}(\widetilde{y}_{n+i}^{(j)}-y_{n+i})^2 \tag{25}
$$

where m is the number of forecast period beyond the end of the sample period; $\widetilde{y}_{n+i}^{(j)}$ is the forecast value for period $n+i$ generated by Model j, $j=1,2$. y_{n+i} is the actually realized value in period $n+i$. In sample theory we have one $\widetilde{y}_{n+i}^{(j)}$. In Bayesian analysis we can generate $\widetilde{y}_{n+i}^{(j)}$ and obtain the distribution of the MSEF given the actually realized y_{n+i} 's. For one period ahead forecast (i.e., $m=1$), the MSEF becomes the squared errors of forecast (SEF):

$$
SEF=(\widetilde{y}_{n+1}^{(j)}-y_{n+1})^2 \tag{26}
$$

1-period ahead PDIC shows CKLS-ARMA-GARCH-NORMAL (i.e., Model

2) has better out-of-sample predictive power since its PDIC value is smaller (see Table 7). 1-period ahead MSEF (i.e. SEF) shows CKLS-ARMA-GARCH-EPD (i.e., Model 1) is better since its MSEF is smaller. The densities and cumulative densities of the 1-period ahead MSEF (i.e., SEF)'s are given in Figure 7, 8, 9 and 10. We see that the cumulative density of the SEF's for Model 1 clearly dominates that for Model 2. That is, CDF of the SEF for model 1 goes to 1 more quickly than that of model 2.

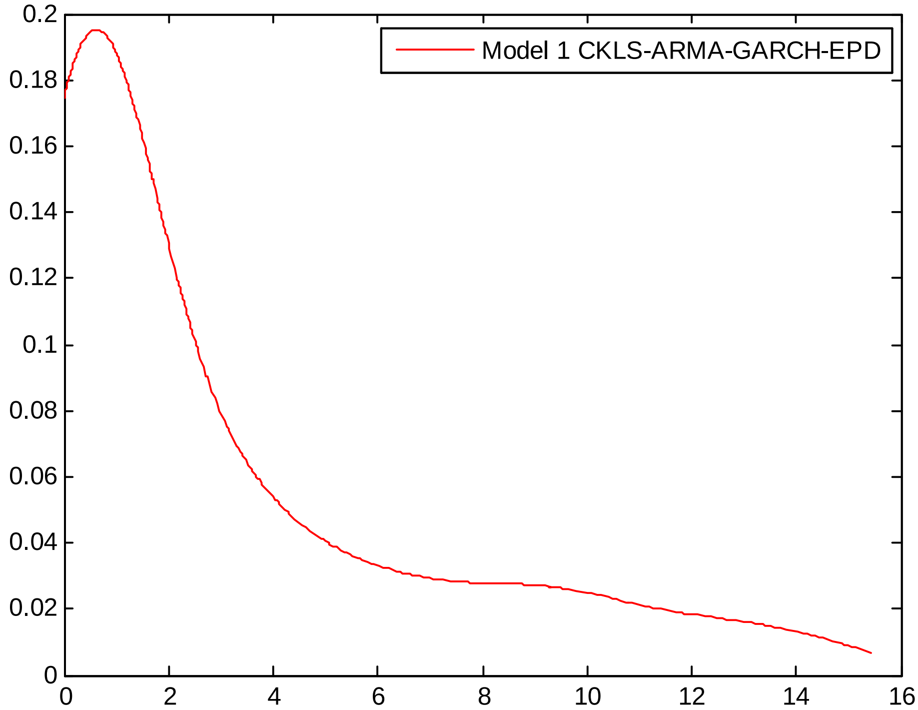

Figure 7: PDF of SEF(Model 1: CKLS-ARMA-GARCH-EPD)

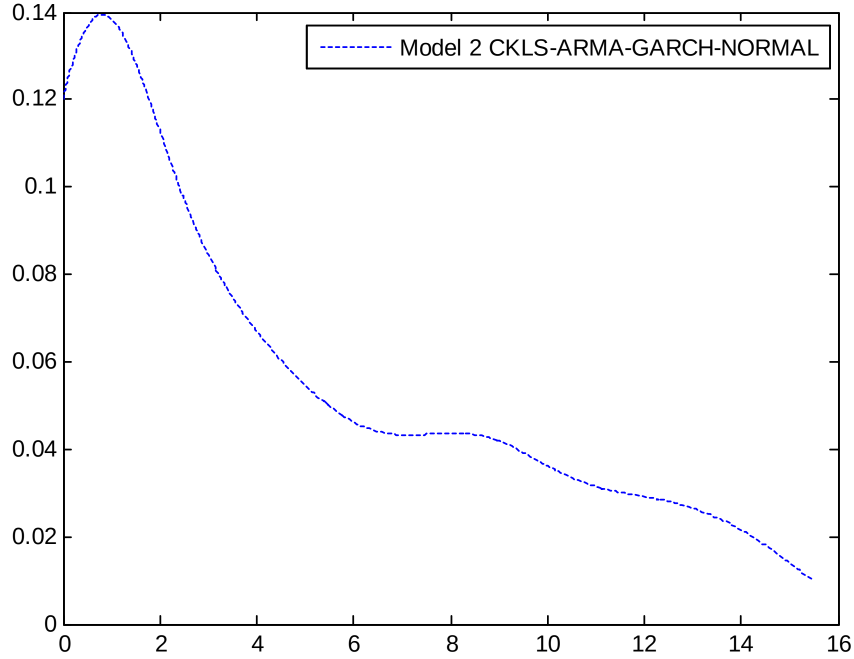

Figure 8: PDF of SEF(Model 2: CKLS-ARMA-GARCH-NORMAL)

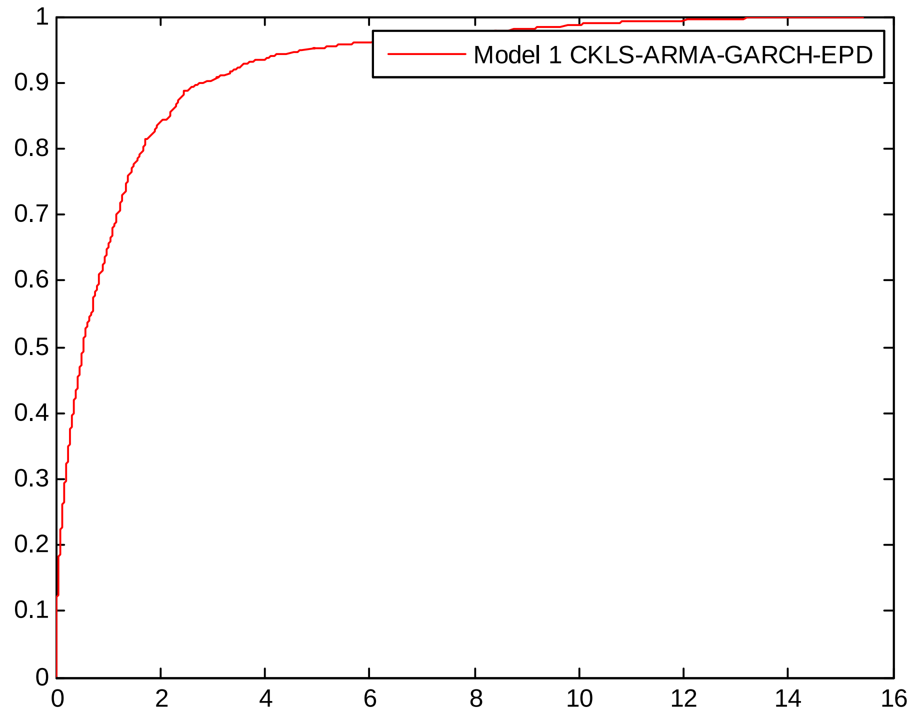

Figure 9: CDF of SEF(Model 1: CKLS-ARMA-GARCH-EPD)

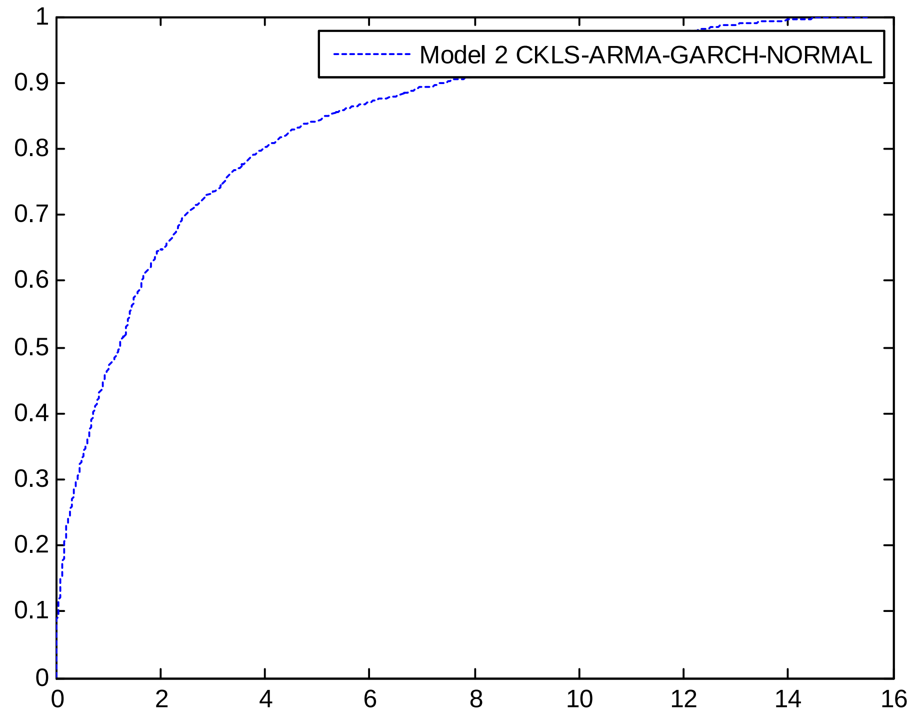

Figure 10: CDF of SEF (Model 2: CKLS-ARMA-GARCH-NORMAL)

2.4.4.1 Check robustness (fixed estimation scheme) To test the robustness of previous results for 1-period ahead forecast, 3-period, 6-period and 9-period ahead values of PDIC and MSEF are also calculated (see Table 7). Table 7 shows the results of 1-period ahead PDIC and MSEF are robust. That is, PDIC shows CKLS-ARMA-GARCH-NORMAL (i.e., Model 2) always has better forecastibility since its PDICs are smaller. And MSEF shows CKLS-ARMA-GARCH-EPD (i.e., Model 1) always has better forecastibility since its MSEFs are smaller. The MSEF of CKLS-ARMA-GARCH-NORMAL increases greatly when out-of-sample period increases. Since MSEF comes from sample theory and PDIC comes from Bayesian theory, we prefer the results coming from PDIC. The results of forecast show the model with better in-sample fit is not necessary the one with better forecastibility.

2.4.4.2 Check robustness (rolling estimation scheme) To test the robustness of PDIC results from fixed estimation scheme, we use rolling estimation scheme to derive PDIC for 10-period, 15-period, 20-period and half-sample ahead data. The results are listed in Table 8. The results of rolling estimation scheme show CKLS-ARMA-GARCH-EPD model has better forecastibility. From Table 8, we can see the PDIC for CKLS-ARMA-GARCH-EPD model becomes smaller as the forecast period increases from 10 to half-sample. For the half sample case, we plot the PDICs of both models in Figure 11. The graph in Figure 11 shows the PDICs of CKLS-ARMA-GARCH-EPD are always smaller than those of CKLS-ARMA-GARCH-NORMAL, and the difference between them is bigger when out-of-sample period increases.

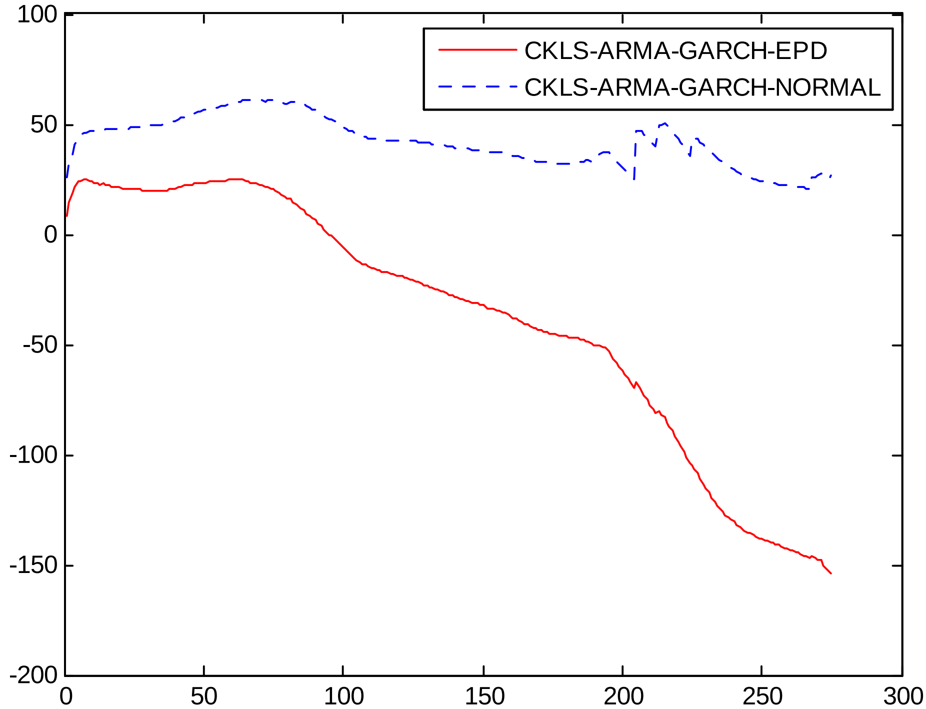

Figure 11: Half-sample PDICs Comparison

2.5 Chapter conclusions

In this paper, we firstly propose a short-term interest rate model. The model is the CKLS-ARMA-GARCH model with error terms distributed as exponential power distribution (EPD). We denote this model as CKLS-ARMA-GARCH-EPD (or Model 1).

	BF21	DIC
Model 1: CKLS-ARMA-GARCH-EPD	7.52	165.76
Model 2: CKLS-ARMA-GARCH-NORMAL	–	245.94

Table 6: In-sample Fit (Jan. 63 - Dec. 07)

	PDIC		MSEF	
	Model 1	Model 2	Model 1	Model 2
1-period ahead	13.07	9.78	3.16(3.91)	4.41(4.33)
3-period ahead	21.98	17.54	4.12(3.92)	5.84(5.97)
6-period ahead	24.11	19.93	3.92(3.73)	35.67(50.93)
9-period ahead	20.31	16.16	3.93(3.74)	145.86(167.46)

Note: Model 1 is CKLS-ARMA-GARCH-EPD.
Model 2 is CKLS-ARMA-GARCH-NORMAL.

Table 7: Multi-period ahead Out-of-Sample Model Selection Using the Estimates from Period Jan. 63 - Dec. 07(Fixed Estimation Scheme)

	PDIC	
	Model 1	Model 2
10-period ahead	-	Smaller
15-period ahead	Smaller	-
20-period ahead	Smaller except the first 10 periods	-
Half sample ahead	Smaller	-

Note: Model 1 is CKLS-ARMA-GARCH-EPD.
Model 2 is CKLS-ARMA-GARCH-NORMAL.

Table 8: Multi-period ahead Out-of-Sample Model Selection (Rolling Estimation Scheme)

Bayesian MCMC algorithms are developed to estimate this model. Then, both in-sample fit and out-of-sample forecastability of this model are compared with CKLS-ARMA-GARCH-NORMAL(or Model 2). Bayes Factor(see Gelfand and Dey (1994)) and Deviance Information Criterion(DIC, see Spiegelhalter(2002)) are used as in-sample fit criteria. Predictive Deviance Information Criterion (PDIC) and Mean Squared Error of Forecast (MSEF) are used as criteria for out-of-sample forecast. Empirical data of US Federal funds rates are analyzed.

Simulation results show that the MCMC algorithms used to estimate Model 1 and 2 are efficient since the filtered Fluctuation test (FT) and filtered Kolmogorov Smirnov test (KS) (see Ploberger, Kramer and Kontrus(1989), Smirnov(1939) and Goldman et al.(2007)) indicate convergence.

We figure out the long-run equilibrium short-term interest rate is 1.53% calculated by the posterior means of CKLS-ARMA-GARCH-EPD. In the paper of CKLS(1992), their long-run equilibrium short-term interest rate is 6.89%. We think the difference is caused by different models, different estimation techniques and different data sets.

The results of in-sample model selection are as follows. DIC shows CKLS-ARMA-GARCH-EPD fits in-sample data better since its DIC value is smaller. However, Bayes Factor has decisive evidence supporting Model 2. DIC uses effective number of parameters instead of the true number of parameters to compare models. So, we prefer the result of DIC criterion and pick CKLS-ARMA-GARCH-EPD model as the one with better in-sample fit.

By fixed estimation scheme, both 1-period and multi-period ahead PDICs show CKLS-ARMA-GARCH-NORMAL model has better out-of-sample forecastibility since its PDIC values are smaller. However, if we use criterion of Mean Squared Error of Forecast(MSEF), then CKLS-ARMA-GARCH-EPD is better. Since MSEF comes from sample theory and PDIC comes from Bayesian theory, we prefer the results coming from PDIC.

Results from both in-sample fit and out-of-sample forecast via fixed estimation scheme show a model with better in-sample fit does not necessary have better out-of-sample forecastibility. Results from out-of-sample forecast via rolling estimation scheme show CKLS-ARMA-GARCH-EPD model can predict long run better since its PDICs are smaller as forecast period increases.

Future extensions include but not limit to follows. We can add macro variables of Bernake et al.(2005) into the CKLS-ARMA-GARCH-EPD model (see Bernake et al.(2005) for the Factor-Augmented (FAVAR) model). One can also introduce a regime switching variable in CKLS-ARMA-GARCH-EPD model (see Dueker et al.(2007)). 2 or 3 latent factors can also be added to CKLS-ARMA-GARCH-EPD model (see Andersen and Lund(1997), Jones(2003), Bali(2000) and Sorwar (2005)). Or one can use our model to price derivatives. Last but not the least, one can check the robustness of our model using international data.

3 Markov switching models for realized volatility

3.1 Motivation

Many financial time series display asymmetric behavior. For example, large negative returns appear more frequently than large positive returns and sudden dramatic political events have big effects on financial time series. In order to capture such asymmetric behavior, nonlinear time series models are developed[8].

A lot of attempts try to model this kind of nonlinearity explicitly. One approach is to define different regimes, i.e., state of the world, and allow the dynamic behavior of economic variable to depend on the regimes (see Priestly (1980, 1988)). That means certain properties of the time series such as its mean, variance or autocorrelation are different in different regimes. In this case, we call this time series state-dependent or regime-switching.

Regime-switching process can be deterministic or stochastic (see Figure 12). One example of deterministic regime process is the seasonal effect. Stochastic regime can be found in LeBaron (1992), Kräger and Kugler (1993)

[8] Various statistical tests have been used to detect the nonlinearity in stock prices and exchange rates (see Hinich and Patterson (1985), Scheinkmann and LeBaron (1989), Hsieh (1989,1991), Crato and de Lima (1994), and Brooks (1996)).

and Chappell et al. (1996).

Stochastic regime-switching models can be divided into two categories (see Figure 12). The first category assumes that the regimes can be characterized or determined by an observable variable, i.e., the regimes that have occurred in the past and present are known with certainty, and can be calculated by statistical tools. Examples are Threshold Autoregressive model (TAR)[9], Self-Exciting TAR (SETAR) model with threshold value determined by the time series itself, or Smooth Transition AR (STAR) model including Logistic STAR (LSTAR)[10].

The second category assumes that the regimes are unobserved or hidden, and they are determined by an underlying unobservable stochastic process, which we denote as s_t . We can not be sure that a particular regime will occur at a particular point of time, but can only assign probabilities to the occurrence of the different regimes. When the property of s_t in a model is specified to be a first order Markov-process, i.e. the current regime only depends on the regime one period ahead , then this kind of model is called Markov switching model(MSW).

Markov switching model (hereafter denoted as MSW) was introduced to economics by Hamilton (1989) . Since then, many extensions of Markov switching model are developed[11]. For example, Cai(1994), and Hamilton and

[9] See Tong (1978), Tong and Lim (1980), and Tong (1990).

[10] See Bacon and Watts (1971), Chan and Tong (1986), Granger and Terävirta (1993), and Teräsvirta (1994, 1998))

[11] See Kim(1994), Cai (1994), Hamilton and Susmel (1994), Gray(1996), Dueker(1996), Kaufman and Fruhwirth-Schnatter (2000), Das and Y00(2004), Yoo(2006), Henneke,

Susmel (1994) extended MSW-AR to MSW-ARCH by adding the Markov switching variable s_t into the ARCH equation. Gray(1996) and Dueker(1997) studied a MSW-GARCH model. Das and Yoo (2004), and Yoo (2006) analyzed MSW-ARMA-GARCH models.

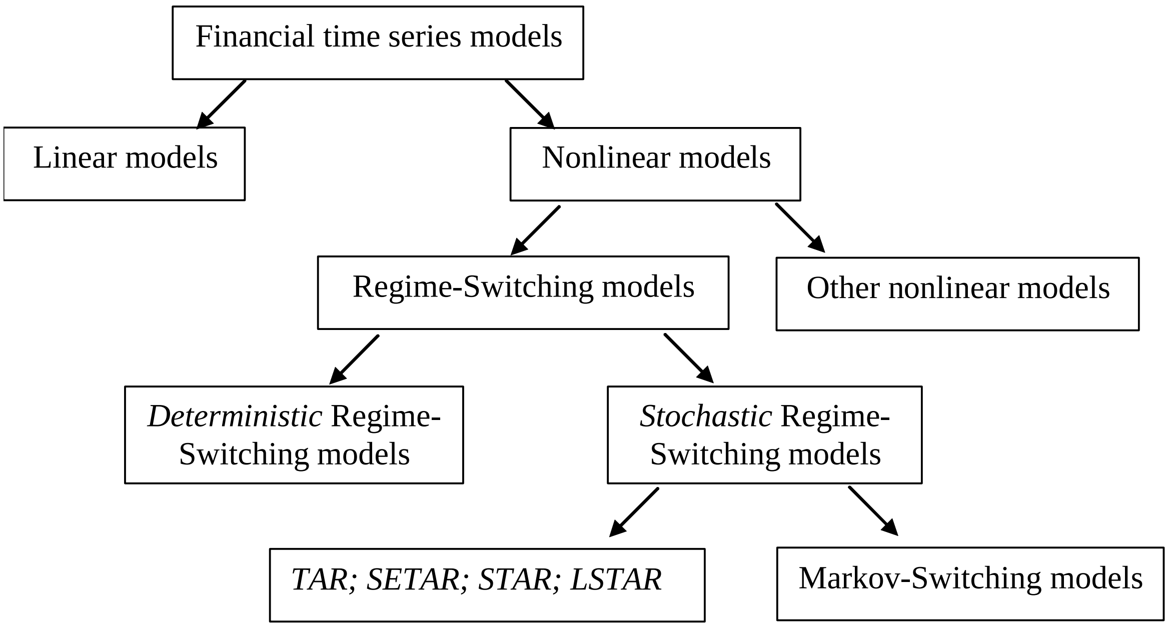

Figure 12: Financial Time Series Models

At the beginning of development of Markov switching models, most of the researches are used maximum likelihood estimation (MLE) to estimate the parameters in the models[12].

However, as the Markov switching models are extended to be more and more complicated, MLE becomes computationally unfeasible. For example,

Rachev and Fabozzi(2006).

[12]See Hamilton (1989), Cai(1994), Hamilton and Susmek (1994), Gray (1996), Dueker(1997), and Klaassen(2002).

when the MA(q) error terms are added into the models, then the likelihood function depends on all previous q periods of the past history of the state variables, which makes the joint process become a non-Markovian process and is hard to be estimated by MLE. In addition, for MSW-GARCH, if there are only 2 states with N sample size, then there will be 2^N cases of the likelihood function to be consider.

There are two ways to deal with previous problem of MLE. One is to use reparameterization to recover a Markovian frame work from the non-Markovian models (see Jones(1987) and Barnett et.al.(1996)). Another is to use Markov Chain Monte Carlo (MCMC) to deal with the non-Markovian structure directly [13] (see Billio et.al.(1999), Kaufmann et.al.(2000), Das and Yoo(2004), Yoo(2006), and Henneke, Rachev and Fabozzi(2006)). MCMC is a very powerful tool for the numerical computation of integrals, which can be used easily to estimate complicated models.

Markov switching models have been widely applied to GDP, T-bill rates, exchange rates and stock prices[14], but there is little application to the realized volatility, a new measure of volatility first introduced by Andersen and Bollerslev (1997). There are a lot of researches done to model and forecast the realized volatility [15] because of the importance of volatility in finance. How-

[13] For more discussion of MCMC, please refer to Robert and Casella (1999).

[14] See Hamilton (1998) for GDP, Cai(1994) and Gray (1996) for T-bill rate, Hamilton and Susmel (1994), and Dueker (1997) for stock prices, Klaassen (2002) for exchange rate.

[15] see Andersen, Bollerslev, Diebold, and Labys(2001, 2003), Andersen, Bollerslev, Diebold and Ebens(2001), Fleming, Kirby and Ostdiek(2003), Maheu and McCurdy(2002), Goldman, Nam and Wang(2005).

ever, most researches use linear models, GARCH, RiskMetrics, FIEGARCH or TARMA.

The first contribution of this paper is to design a random walk draw algorithm for a Markov switching model (MSW-ARMA-GARCH-S_t-in-variance). And then the estimates from this algorithm are compared with those of non-random walk draw (see Das and Yoo(2004)).

The 2nd contribution of this paper is to compare two Markov switching models by MSE and MSEF. We try to test one hypothesis: is the model with s_t in variance better than the one with s_t in mean? A Markov-switching model with state variable in variance (i.e., MSW-ARMA-GARCH-S_t-in-variance) is estimated by MCMC algorithm of Das and Yoo(2004). And another Markov-switching model with state variable in mean (i.e., MSW-ARMA-GARCH-S_t-in-mean) is estimated by MCMC algorithm of Yoo(2006). Metropolis Hasting algorithm and Gibbs samplers are used. In order to generate latent variable s_t, single-move method [16] is used.

The 3rd contribution is to study the realized volatility using Markov switching models. Even though there is few research on realized volatility using Markov-switching models, these MSW models are different from ours. For example, Maheu and McCurdy(2002) found that a duration dependent Markov-switching ARMA model is better than other linear models for the realized volatility of the DM/$ exchange rate. Different from us, their model

[16] For single-move method, see Carlin, Polson, and Stoffer (1992), and Albert and Chib (1993). For multi-move, see Carter and Kohn (1994),and Kim and Nelson (1999).

has no GARCH term.

This paper is organized as follows. Section 2 provides details about the Markov switching ARMA-GARCH models. Bayesian inference and the MCMC algorithm for the estimation are given in section 3 and 4. Section 5 explains the effectiveness of the MCMC algorithm using numerical examples. In section 6, the estimation results from the realized volatility of MMM company are presented. Conclusions are given in section 7.

3.2 Markov switching models

The Markov switching models with ARMA-GARCH errors are as follows:

3.2.1 Model 1: MSW-ARMA-GARCH-S$_t$-in-variance

$$y_t = x_t\gamma + u_t, \tag{27}$$

$$u_t = \sum_{i=1}^{p} \phi_i u_{t-i} + e_t + \sum_{i=1}^{q} \theta_i e_{t-i}, \quad e_t \sim N(0, \sigma_t^2), \tag{28}$$

$$\sigma_t^2 = \alpha_0 + s_t\mu + \sum_{i=1}^{r} \alpha_i e_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}^2, \tag{29}$$

$$\alpha_0 > 0, \alpha_0 + \mu > 0, \alpha_i \geq 0, i = 1, 2, ...r, \beta_i \geq 0, i = 1, 2, ...s,$$

$$1 \geq \sum_{i=1}^{max(r,s)} (\alpha_i + \beta_i) \tag{30}$$

For model 1, the Markov switching variable s_t is in conditional variance (i.e., GARCH term, equation 29).

3.2.2 Model 2: MSW-ARMA-GARCH-S$_t$-in-mean

$$y_t = x_t\gamma_1 + s_t\mu + u_t, \tag{31}$$

$$u_t = \sum_{i=1}^{p} \phi_i u_{t-i} + e_t + \sum_{i=1}^{q} \theta_i e_{t-i}, \quad e_t \sim N(0, \sigma_t^2), \tag{32}$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{r} \alpha_i e_{t-i}^2 + \sum_{i=1}^{s} \beta_i \sigma_{t-i}^2, \tag{33}$$

$$\alpha_0 > 0, \alpha_i \geq 0, i = 1, 2, ...r, \beta_i \geq 0, i = 1, 2, ...s,$$

$$\sum_{i=1}^{max(r,s)} (\alpha_i + \beta_i) \leq 1. \tag{34}$$

For model 2, the Markov switching variable s_t is in the mean component of dependent variable (equation 31).

For both models, we assume the state variable s_t is a two- state, first order Markov switching process. The hidden transition probabilities for this Markov switching process are :

$$Pr[s_t = 0|s_{t-1} = 0] = p_{00}, Pr[s_t = 1|s_{t-1} = 0] = p_{01}. \tag{35}$$

$$Pr[s_t = 0|s_{t-1} = 1] = p_{10}, Pr[s_t = 1|s_{t-1} = 1] = p_{11}$$

We also impose constraints (30) and (34) on the coefficients of GARCH to guarantee that the conditional variance is positive.

For notation simplicity, we use following parameter definitions for model

1:

$$\phi = [\phi_1, ...\phi_p]', \quad \theta = [\theta_1, ...\theta_q]', \quad \alpha = [\alpha_0, \mu, \alpha_1, ...\alpha_r]', \beta = [\beta_1, ..., \beta_s]',$$

$$\delta = [\gamma, \phi, \theta, \alpha, \beta, p_{00}, p_{11}]', \quad S = [s_1, ..., s_T]', \quad Y = [y_1, ..., y_T]'.$$

For model 2, in addition to previous definitions, we define two more notations: $\alpha = [\alpha_0, \alpha_1, ...\alpha_r]', \quad \gamma = [\gamma_1, \mu]'$.

3.3 Bayesian inference

For Bayesian inference[17], we need to derive the posterior distribution of all parameters and the state variables $S = [s_1, ..., s_T]'$ conditional on the observed data Y. According to Bayes' rule, the posterior distribution of model 1 or model 2 is:

$$p(\delta, S|Y) = p(\delta, S, Y)/p(Y) \quad \text{(property of conditional density)} \tag{36}$$

$$= p(\delta, S)p(Y|\delta, S)/p(Y) \quad \text{(property of conditional density)} \tag{37}$$

$$\propto p(\delta, S)p(Y|\delta, S) \quad (p(Y) \text{ is known}) \tag{38}$$

$$\propto p(\delta)p(S|\delta)p(Y|\delta, S) \quad \text{(property of conditional density)} \tag{39}$$

The first term of equation (39) is the prior distribution for all parameters.

[17]In the classical framework, inference on Markov-switching models first estimates the models's unknown parameters, then makes inferences on the unobserved Markov switching variables $s_t, t = 1, ..., T$. In the Bayesian analysis, both the parameters and the Markov-switching variable $s_t, t = 1, ..., T$ are treated as random variables.

This prior $p(\delta)$is assumed to be:

$$\begin{aligned} p(\delta) &= p(\gamma, \phi, \theta, \alpha, \beta, p_{00}, p_{11}) && (40)\\ &= p(\gamma)p(\phi)p(\theta)p(\alpha)p(\beta)p(p_{00})p(p_{11}) && (41)\\ &= N(\mu_\gamma, \Sigma_\gamma) * N(\mu_\phi, \Sigma_\phi) * N(\mu_\theta, \Sigma_\theta) && (42)\\ &\quad * N(\mu_\alpha, \Sigma_\alpha) * N(\mu_\beta, \Sigma_\beta)\\ &\quad * Beta(u_{00}, u_{01}) * Beta(u_{11}, u_{10}) \end{aligned}$$

The equal sign in (41) comes from the assumption of independence, i.e., we assume the prior distributions of parameters are independent. $N(.)$ is the normal density function, and $Beta(.)$ is the beta density function.

The second term of equation (39) is :

$$\begin{aligned} p(S|\delta) &= p(S|\gamma, \phi, \theta, \alpha, \beta, p_{00}, p_{11}) && (43)\\ &= p(S|p_{00}, p_{11}) \quad \text{Since the independence assumption.}\\ &= \prod_{t=1}^{T} p(s_{t+1}|s_t, p_{00}, p_{11}) \quad \text{Since Markov property.}\\ &= p_{00}^{n_{00}}(1-p_{00})^{n_{01}} p_{11}^{n_{11}}(1-p_{11})^{n_{10}} \end{aligned}$$

where n_{ij} is the number of the transitions from state i to state j.

The third term in equation (39) is :

$$p(Y|\delta, S) = \prod_{t=1}^{T-1} p(y_t|y_{t-1}, \delta, S) \tag{44}$$
$$= \prod_{t=1}^{T-1} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp[-\frac{e_t^2}{2\sigma_t^2}]$$

3.4 MCMC implementation

Markov Chain Monte Carlo(MCMC) algorithm is used to estimate the parameters. MCMC is one way of the numerical integrations based on the Clifford-Hammersley theorem. This theorem shows that a joint distribution can be characterized by its complete conditional distributions. When applied in Bayesian, that means the posterior distributions, $p(\delta, S|Y)$, can be characterized by its complete conditional distributions: $p(\delta|S, Y)$ and $p(S|\delta, Y)$.

In the MCMC algorithm, both Gibbs Samplers (GS) and Metropolis Hasting (MH) are used. Gibbs Samplers are used to draw p_{00}, p_{11} since its complete conditional distribution is assumed to be beta distribution, which is known. Metropolis Hasting (MH) algorithms are used to draw other parameters such as $\gamma, \phi, \theta, \alpha,$ and β, since the complete conditional distributions of these parameters are unknown.

3.4.1 Algorithm using non-random walk draw

We draw parameters such as $\gamma, \phi, \theta, \alpha,$ and β using non-random walk draw(see Appendix i), the probability to accept a proposal value is

$$\rho = min\{\frac{p(\hat{\delta}|Y,S) * f(\delta^{(j-1)})}{p(\delta^{(j-1)}|Y,S) * f(\hat{\delta})}, 1\} \tag{45}$$

Where $f(.)$ is the proposal distribution function. δ^{j-1} is the $j-1$ draw. $\hat{\delta}$ is a generated value for δ. We use the method of Nakatsuma(2000) to choose the proposal distribution $f(.)$[18].

3.4.2 Algorithm using random walk draw

We also use random walk draw in the MCMC algorithm to draw parameters $\gamma, \phi, \theta, \alpha,$ and β for model 1 (see Appendix ii). In the random walk draw, the probability to accept a proposal value is $\rho = min\{\frac{p(\hat{\delta}|Y,S)}{p(\delta^{(j-1)}|Y,S)}, 1\}$, since the proposal distribution function $f(.)$ in equation 45 is symmetric, i.e.,$f(\delta^{(j-1)}, \hat{\delta}) = f(\hat{\delta}, \delta^{(j-1)})$.

Since $S = [s_1, ..., s_T]$ are latent variables, data augmentation are used. That is, the paramenter δ is augmented with the states S. Single move method is used to generate S since the process is non-Markovian and multi-move method is infeasible.

Parameters in δ are divided into 3 groups: $\delta_1 = (p_{00}, p_{11})$, $\delta_2 = (\gamma, \phi, \theta)$, $\delta_3 =$

[18] Almost any distribution can be chosen as the proposal distribution. However, for the speed of the convergence, it is crucial to select a proper proposal distribution. See Robert and Casella(1999).

(α, β). Then, the MCMC algorithm is :

1. Draw s_t from $p(s_t|S_{-t}, Y, \delta)$ for all $t = 1, 2, ...T$ by single-move GS method.

2. Draw δ_1 from $p(\delta_1|S, Y, \delta_2, \delta_3) = p(\delta_1|S)$, the Beta distribution, by GS.

3. Draw δ_2 from $p(\delta_2|S, Y, \delta_3, \delta_1) = p(\delta_2|S, Y, \delta_3)$ by MH algorithm.

4. Draw δ_3 from $p(\delta_3|S, Y, \delta_2, \delta_1) = p(\delta_3|S, Y, \delta_2)$ by MH algorithm.

We use equations (27), (28) for model 1 (or equations (31), (32) for model 2) to draw proposal density for $\delta_2 = (\gamma, \phi, \theta)$. The corresponding likelihood function is:

$$p(Y|S, \delta_2, \delta_3) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp[-\frac{e_t^2}{2\sigma_t^2}] \tag{46}$$

For $\delta_3 = (\alpha, \beta)$, we follow the approximation of Nakatsuma (2000) to get proposal densities:

$$e_t^2 = \alpha_0 + \sum_{i=1}^{l} (\alpha_i + \beta_i) e_{t-i}^2 + w_t - \sum_{i=1}^{s} \beta_i w_{t-j}, \quad w_t \sim N(0, 2\sigma_t^2), \tag{47}$$

$$\text{where:} \quad l = max\{r, s\}, \quad \alpha_i = 0 \quad for \quad i > r, \quad \beta_i = 0 \quad for \quad i > s,$$

$$e_t^2 = 0, \quad w_t = 0 \quad \text{for} \quad t \leq 0$$

The corresponding likelihood function is :

$$p(e^2|Y, S, \delta_2, \delta_3) = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi(2\sigma_t^4)}} \exp[-\frac{w_t^2}{2(2\sigma_t^4)}] \quad (48)$$

$$\text{where} \quad e^2 = [e_1^2, ..., e_T^2]'.$$

For more detailed of MCMC implementation, please refer to the Appendix i & ii.

3.5 Simulation results

In this section, we use simulation examples to show the effectiveness of the MCMC algorithms. We estimate model 1 and model 2 while setting $p = 1, q = 1, r = 1, s = 1$, i.e., two Markov switching models with ARMA(1,1) -GARCH(1,1) error term.

We first design an algorithm using random walk draws to estimate model 1 (see Appendix ii). The results[19] are listed in Table 9. The sample size is 1000, the draws discarded are 1000 and every 5th draw is kept. It takes about 6 hours to run the program. In Table 9, all parameters are close to true value except μ, α_1 and p_{11}. The parameter μ has an opposite sign compared with the true parameter. The mean of α_1 is double of the true parameter. For p_{11}, the value is just a little bit higher than 0.5, which is far away from true value 0.98.

[19] We try different parameters and sample size to see the results.

Then, we estimate model 1 using non-random walk draw with the proposal densities listed in Appendix i. The results for model 1 are in Table 10. Sample size is 500, 100 draws are discarded and every 1 draw of 500 is used. It takes 44 minutes to get the result. In Table 10, every posterior mean is closed to its true value. Compared with Table 9, the results are obvious better than that. So we choose this algorithm for model 1 to estimate the realized volatility in next section[20].

The simulation results for model 2 are in Table 11[21]. Sample size is 200, draws discarded are 8000 and every 2 draws from 3000 are used. It takes one and a half hour to get the result. In Table 11, every posterior mean is closed to its true value.

In Figure 13 and 14, we can see that the estimated state variable $\hat{s}_t$ from the simulations are very similar to the true state variable s_t. So, the MCMC algorithm in Appendix i estimates both the parameters and the states very well.

[20]Note: MSW-ARMA-GARCH-S_t-in-variance is estimated by MCMC algorithm of Das and Yoo(2004).

[21]Note: MSW-ARMA-GARCH-S_t-in-mean is estimated by MCMC algorithm of Yoo(2006).

	True	Mean	St.Dev.
γ_1	0	$-.036$	.013
μ	.05	$-.407$	2.724
ϕ_1	.7	.673	.026
θ_1	.3	.313	.045
α_0	.001	.001	.000
α_1	.3	.615	.073
β_1	.5	.532	.034
P_{00}	.98	.997	.002
P_{11}	.95	.562	.179
$\alpha_1 + \beta_1$	.80	1.148	.049

Table 9: Simulation Results for Model 1 (Random Walk Draw)

	True	Mean	St.Dev.
γ_1	0	.014	.066
μ	.05	.074	.024
ϕ_1	.70	.764	.046
θ_1	.30	.248	.074
α_0	.01	.037	.039
α_1	.30	.291	.067
β_1	.50	.469	.113
P_{00}	.98	.976	.019
P_{11}	.95	.960	.032
$\alpha_1 + \beta_1$	.80	.760	.089

Table 10: Simulation Results for Model 1 (Non-Random Walk Draw)

	True	Mean	St.Dev.
γ_1	0	.015	.027
μ	.4	.408	.034
ϕ_1	.7	.571	.084
θ_1	.2	.307	.096
α_0	.005	.007	.002
α_1	.4	.300	.133
β_1	.3	.238	.158
P_{00}	.95	.922	.025
P_{11}	.90	.858	.046
$\alpha_1 + \beta_1$	.70	.539	.157

Table 11: Simulation Results for Model 2 (Non-random Walk Draw)

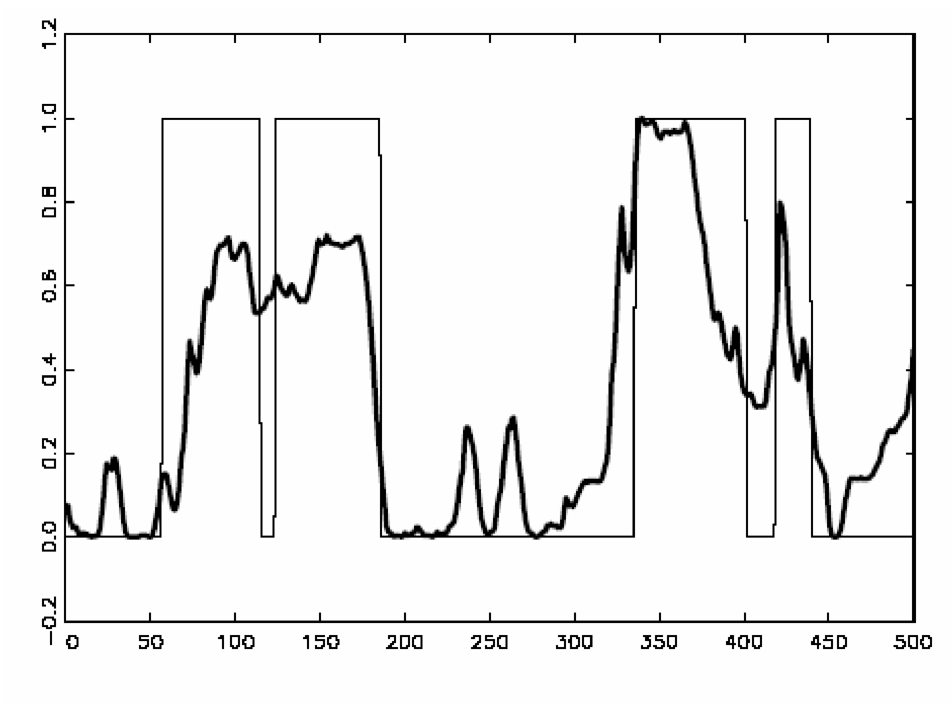

Figure 13: S(t) and Estimated S(t) for Model 1 (Note: The bold line is the estimated value.)

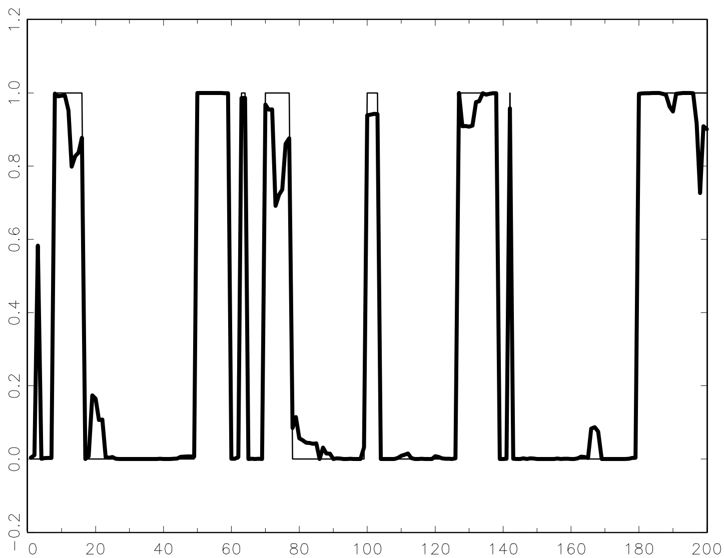

Figure 14: S(t) and Estimated S(t) for Model 2 (Note: The bold line is the estimated value.)

3.6 Application to realized volatility

In this section, we apply the algorithms for model 1 and model 2 to the realized volatility of MMM company. The data are 5-minute returns from TAQ database. The sample period is from January 4, 1993 to December 31, 2004. We use 600 observations. The first 500 data are used for estimation and in-sample fit evaluation. And the last 100 observations are used for out-of-sample forecast evaluation.

The results are listed in Table 12. From these estimates, it is hard to tell which model is better. So we calculate in-sample Mean Squared Errors (MSE) and Mean Squared Error of Forecast (MSEF) for out-of sample forecast evaluation (see Table 13). Model 2 has better in-sample fit since its MSE value is smaller. But model 1 has better out-of-sample forecastability since it has lower MSEF.

3.7 Chapter conclusions

Markov switching models have been widely applied to GDP, T-bill rates, exchange rates and stock prices[22]. In this paper, we design a random walk algorithm for a Markov switching model with regime switching in variance (i.e., MSW-ARMA-GARCH-S_t-in-variance). Then, the estimates from this algorithm are compared with those of Das and Yoo(2004).

Also, we compare two Markov switching models by MSE and MSEF. A Markov-switching model with state variable in variance (i.e., MSW-ARMA-GARCH-S_t-in-variance) is estimated by MCMC algorithm of Das and Yoo(2004). And another Markov-switching model with state variable in mean (i.e., MSW-ARMA-GARCH-S_t-in-mean) is estimated by MCMC algorithm of Yoo(2006). Metropolis Hasting algorithm and Gibbs Samplers are used. In order to generate latent variable s_t, single-move method[23] is used.

Lastly, we use previous two Markov switching models to analyze the realized volatility of MMM. The realized volatility is a new measure of volatil-

[22] See Hamilton (1998) for GDP case. See Cai(1994) and Gray (1996) for T-bill rate case, Hamilton and Susmel (1994), and Dueker (1997) for stock prices cases, Klaassen (2002) for exchange rate case.

[23] For single-move method, see Carlin, Polson, and Stoffer (1992), and Albert and Chib (1993). For multi-move, see Carter and Kohn (1994),and Kim and Nelson (1999).

	Model 1		Model 2	
	Mean	St.Dev.	Mean	St.Dev.
γ_1	$-.0006$	.022	.041	.002
μ	.981	.008	.093	.223
ϕ_1	$-.115$	.000	.984	.008
θ_1	0	.000	$-.249$	.000
α_0	.411	.000	.000	.000
α_1	.033	.000	.525	.088
β_1	.084	.000	.652	.034
P_{00}	.997	.002	.994	.004
P_{11}	.897	.086	.244	.196
$\alpha_1 + \beta_1$	.444	.000	1.177	.067

Note:Model 1: MSW-ARMA-GARCH-S_t-in-variance
Model 2: MSW-ARMA-GARCH-S_t-in-mean

Table 12: Estimation Results for the Realized Volatility of MMM Company

Models	**MSE**	**MSEF**
Model 1: MSW-ARMA-GARCH-S_t-in-variance	.0019	.0006
Model 2: MSW-ARMA-GARCH-S_t-in-mean	.0009	.9678

Table 13: MSE and MSEF for the Realized Volatility of MMM Company

ity, first introduced by Andersen and Bollerslev (1997). There are a lot of researches being done to model and forecast the realized volatility. Most of these researches are analyzed the realized volatility by linear models, GARCH, RiskMetrics, FIEGARCH or TARMA models. However, using MSW-ARMA-GARCH models to model the realized volatility is new.

The conclusions of this paper are as follows. Algorithm of random walk draw for model 1 is not better than the algorithm of non-random walk draw in Das and Yoo(2004). According to MSE, we find out model 2 has better in-sample fit. However, MSEF shows model 1 has better forecastibility.

Future extension could be follows. First, we can compare Markov switching model with other nonlinear models such as TARMA. Secondly, we can consider other model selection criteria. Thirdly, we can also extend previous Markov switching models by assuming other error distributions such as the exponential power distribution. Last but not least, we can design a new model which includes two regime switching variables: one regime switching variable in the mean component, another regime switching variable in the variance component.

4 Bayesian analysis of SP500 and FTSE100 using a Gaussian Copula model

4.1 Motivation

In this paper, we first design a Gaussian copula model with ARMA-GARCH-EPD error terms (i.e., Copula-ARMA-GARCH-EPD). Then, we design Markov Chain Monte Carlo (MCMC) algorithms to estimate this model. Monthly data on SP500 and FTSE100 are analyzed.

Copulas (see Sklar(1959)) are convenient tools to construct multivariate joint distributions from margins. In financial risk assessment and actuarial analysis, dependence modeling with copula functions is widely used. Li (2000) introduces Gaussian copula model into pricing Collateralized Debt Obligations (CDOs).

Researches on copulas have been fast growing for the following reasons: (1) As shown in Blyth (1996), Shaw (1997), and Embrechts *et.al.* (1999), copulas can handle nonlinear dependence measures such as Kendall's τ, Spearman's ρ, and Gini indice γ that may be more appropriate for measuring dependence among data.[24] (2) The empirical distributions of many financial time series data are skewed and leptokurtic. This implies that models with normal error distribution may be inadequate. Copulas allow us to build multivariate distributions from leptokurtic and/or skewed marginal distributions,

[24] Additional references are Karolyi and Stultz (1996), Forbes and Rigobon (1999), Straetmans (2000), Login and Solnik (2001), and Ang and Bekaert (2002).

and thus we can find dependence among data.[25]

Papers applying copulas are many and the application of copulas spreads across various fields. For example, in health economics, Zimmer and Trivedi (2006) applied a trivariate copula to analyze family health care demand. Romeo *et.al.* (2006) used bivariate copula models to study multivariate survival data. Li *et.al.* (2006) used a copula variance-components method for genetic studies.[26] In finance, Meucci (2006) used copula-opinion pooling methodology to extend the Black-Litterman model for portfolio management. Arakelian and Dellaportas (2006) studied a threshold model using copulas. Patton (2006) used copulas to model five Asian stock indices to measure financial contagion.[27]

Bayesian and maximum likelihood procedures have been used to estimate copula models. For example, Xu (2004) designed an MCMC algorithm to estimate a mixed copula model. Chen *et.al.* (2004) developed two goodness-of-fit tests for copula models. Chen *et.al.* (2005) proposed a sieve maximum likelihood estimation (MLE) to efficiently estimate copula models. An efficient Bayesian approach for estimating a Gaussian copula model was presented by Pitt, Chan and Kohn (2006).

In this paper we combine the Gaussian copula with ARMA-GARCH-EPD

[25] See Mandelbrot (1963), Fama (1965), Praetz (1972), Blattberg and Gonedes (1974).

[26] See Shih and Louis (1995), Andersen (2004, 2005), Cameron *et.al.* (2004), Andersen *et.al.* (2005), Trivedi and Zimmer (2006) for more copula application in health economics.

[27] See Li (2000), Longin and Solnik (2001), Dias and Embrechts (2004), Rodriguez (2003), Ausin and Lopes (2006), Hu (2006), Bartram *et.al.* (2006), Patton (2006a), and Jondeau and Rockinger (2006) for more copula application in finance.

marginal distributions. Results show the MCMC algorithm is convergent by the filtered Fluctuation test and Kolmogorov-Smirnov test in Goldman *et.al.*(2007). Using monthly data on SP500 and FTSE100, we estimate the Gaussian copula model with ARMA-GARCH-EPD margins.

The organization of the paper is as follows. Section 2 discusses the model: Gaussian copula with ARMA-GARCH-EPD margins. Section 3 presents the posterior density for Copula-ARMA-GARCH-EPD. MCMC algorithm and a numerical example are showed in section 4. Section 5 illustrates an empirical analysis of SP500 and FTSE100. Concluding remarks are given in section 6.

4.2 Copula-ARMA-GARCH-EPD model

The Copula-ARMA-GARCH-EPD is a multivariate regression model:

$$Y = Xb + u \tag{49}$$

where $Y = \begin{bmatrix} Y_1 \\ ... \\ Y_M \end{bmatrix}, X = \begin{bmatrix} X_1 & 0 & ... & 0 \\ 0 & X_2 & .. & 0 \\ & ... & ... & \\ 0 & 0 & ... & X_M \end{bmatrix}, b = \begin{bmatrix} b_1 \\ ... \\ b_M \end{bmatrix}, u = \begin{bmatrix} u_1 \\ ... \\ u_M \end{bmatrix}.$

Or,

$$Y_i = X_i b_i + u_i, \; i = 1, ..., M. \tag{50}$$

where

$$
\begin{aligned}
u_{it} &= \sum_{j=1}^{p} \phi_{ij} u_{it-j} + \sum_{j=1}^{q} \theta_{ij} e_{it-j} + e_{it} &(51)\\
e_{it} &= \sigma_{it}\varepsilon_{it},\ \varepsilon_{it} \sim EPD(1,\ \zeta_i) &(52)\\
\sigma_{it}^2 &= \alpha_{i0} + \sum_{j=1}^{r} \alpha_{ij} e_{it-j}^2 + \sum_{j=1}^{s} \beta_{ij} \sigma_{it-j}^2 &(53)\\
\alpha_{i0} &> 0,\ \alpha_{ij} \geq 0,\ j = 1, ..., r,\ \beta_j \geq 0,\ j = 1, ..., s. &(54)\\
1 &\geq \sum_{j=1}^{\max(r,s)} (\alpha_j + \beta_j) &(55)
\end{aligned}
$$

$Y_i = (y_{1i}, ..., y_{Ti})'$ is a $T \times 1$ vector, $X_i = (x_{1i}^{'}, ..., x_{Ti}^{'})'$ is a $T \times k_i$ matrix of regressors, b_i is a $k_i \times 1$ vector of regression coefficients, and u_i is a $T \times 1$ vector of error terms. The sample size is T and M is the number of equations. And $f_i(\varepsilon_{it})$ is the PDF of exponential power distribution (EPD)[28]:

$$
f_i(\varepsilon_{it}) = \frac{1}{2^{1+1/\zeta_i}\sigma_{it}\ \Gamma(1+1/\zeta_i)} \exp\left\{-\frac{1}{2}\left|\frac{\varepsilon_{it}}{\sigma_{it}}\right|^{\zeta_i}\right\} \tag{56}
$$

$$
0 < \sigma_{it} < \infty,\ 0 < \zeta_i < \infty,\ i = 1, ..., M.
$$

$EPD(\sigma_{it}, \zeta_i)$ becomes the normal distribution when $\zeta_i = 2$:

$$
f_i(\varepsilon_{it}) = \frac{1}{\sqrt{2\pi}\sigma_{it}} \exp\left\{-\frac{1}{2}\left|\frac{\varepsilon_{it}}{\sigma_{it}}\right|^{2}\right\},\ 0 < \sigma_{it} < \infty,\ i = 1, ..., M. \tag{57}
$$

[28] For density of EPD , please refer to Tsurumi and Shimizu(2006).

4.3 Posterior density

4.3.1 Joint density through copula

Let $X_1, ..., X_M$ be the random variables. Their joint density through copula[29] is given by

$$h(X_1, ..., X_M) = c(F_1(x_1), ..., F_M(x_M); \rho) \prod_{i=1}^{M} f_i(X_i) \tag{58}$$

where $c(F_1(x_1), ..., F_M(x_M); \rho)$ is the density of copula; $F_i(x_i)$ is the cumulative function and $F_i(x_i) = \int_{-\infty}^{x_i} f_i(X_i)dX_i$; $f_i(X_i)$ is the marginal density of random variable X_i; x_i is the realized value for random variable X_i, $i = 1, .., M$.

If all the marginal distributions $f_i(X_i)$'s are normal, then the joint density in equation 58 is the multivariate normal PDF. If X_i is not normal, then the joint density $h(X_1, ..., X_M)$ is not multivariate normal even if we use the Gaussian copula (see Appendix iv). We can construct a non-normal joint distribution from margins, and then we can estimate the correlation matrix ρ.

[29] For definition of copula, please see Sklar's Theorem on page 18 of Nelsen (2005).

4.3.2 Gaussian copula density

The density of Gaussian copula[30] is given by

$$c(F_1(e_1),...,F_M(e_M);\rho) = |\rho|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}A'(\rho^{-1}-I)A\right\} \tag{59}$$

where

$$A = (a_1,...,a_M)' \tag{60}$$

$$a_i = \Phi^{-1}(F_i(e_i)) \tag{61}$$

$$F_i(e_i) = \int_{-\infty}^{e_i} f_i(\varepsilon_i)d\varepsilon_i \tag{62}$$

And $f_i(\varepsilon_i)$ is the PDF of ε_i. $F_i(e_i)$ is the CDF of ε_i. ρ is a positive definite matrix with diagonal elements being unity. $\Phi(.)$is the standardized normal CDF. And I is the identity matrix.

4.3.3 Posterior density for Copula-ARMA-GARCH-EPD

We define $\Theta_i \equiv \{b_i, \{\phi_{ij}\}_{j=1...p}, \{\theta_{ij}\}_{j=1...q}, \{\alpha_{ij}\}_{j=0...r}, \{\beta_{ij}\}_{j=1...s}, \zeta_i\}$ and $\Theta \equiv (\Theta_1,...,\Theta_M,\rho)$. The posterior density of Θ for Copula-ARMA-GARCH-EPD is

$$p(\Theta|Y,X)=\prod_{t=1}^{T} c(F_1(e_{1t})...F_M(e_{Mt});\rho)f_1(\varepsilon_{1t})...f_M(\varepsilon_{Mt}) * p(\Theta)$$

[30] For density of copula , please refer to page 8 of Bouyé (2000).

where

$$
\begin{aligned}
c(F_1(e_{1t})...F_M(e_{Mt});\rho) &= |\rho|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}A_t'(\rho^{-1}-I)A_t\right\} && (63)\\
A_t &= (a_{1t},...,a_{Mt})' && (64)\\
a_{it} &= \Phi^{-1}(F_{it}(e_{it})) && (65)\\
F_{it}(e_{it}) &= \int_{-\infty}^{e_{it}} f_i(\varepsilon_{it})d\varepsilon_{it} && (66)\\
f_i(\varepsilon_{it}) &= \frac{1}{2^{1+1/\zeta_i}\sigma_{it}\ \Gamma(1+1/\zeta_i)} \exp\left\{-\frac{1}{2}|\frac{\varepsilon_{it}}{\sigma_{it}}|^{\zeta_i}\right\} && (67)\\
0 &< \sigma_{it}<\infty,\ 0<\zeta_i<\infty,\ i=1,...,M. && (68)\\
\varepsilon_{it} &= e_{it}/\sigma_{it} && (69)\\
\sigma_{it}^2 &= \alpha_{i0}+\sum_{j=1}^{r}\alpha_{ij}e_{it-j}^2+\sum_{j=1}^{s}\beta_{ij}\sigma_{it-j}^2 && (70)\\
e_{it} &= u_{it}-\sum_{j=1}^{p}\phi_{ij}u_{it-j}-\sum_{j=1}^{q}\theta_{ij}e_{it-j} && (71)\\
u_{it} &= Y_{it}-X_{it}b_{it}, \quad i=1,...,M. \quad t=1,...,T. && (72)
\end{aligned}
$$

$p(\Theta)$ is the prior density for Θ. We use prior: $p(\Theta)=1$.

4.4 MCMC algorithms and simulation results

We estimate the parameters θ of Copula-ARMA-GARCH-EPD by the MCMC algorithms that are explained in the Appendix iii. The sample size is 300. We run 18000 MCMC iterations and discard the first 3000 draws. Every 10th draw is kept. All acceptance rates are higher than .31. It takes 16 hours to get the results.

Parameters	P-values of FT	P-values of KS
γ_1	.28	1
ϕ_1	1	.87
θ_1	1	.76
ϖ_1	.15	.00
α_1	.99	.01
β_1	.59	.00
ζ_1	.96	.05
γ_2	.44	.63
ϕ_2	1	.21
θ_2	1	.72
ϖ_2	.28	.00
α_2	.92	.59
β_2	.88	.02
ζ_2	1	.39
ρ	1	.54

Table 14: P-values of Filtered Fluctuation Tests and Filtered Komogorov-Smirnov Tests

We check the convergence of the draws by plotting the draws and by the filtered Fluctuation test and Kolmogorov-Smirnov test that are given in Goldman *et.al.*(2007). The plots of the draws indicate convergence (see Figure 15 and 16). And the convergence tests also indicate convergence since the p-value of filtered Fluctuation test or filtered Kolmogorov-Smirnov test is no less than 5% significance level[31](see Table 14).

The summary statistics of the MCMC draws are given in Table 15. The

[31] We judge convergence by 3 criteria: 1) eye rolling convergence. i.e., the draws plotted in a graph are convergent, 2) P-value of Fluctuation test is no less than 5% significance level, 3) or P-value of Kolmogrov-Smirnov test is no less than 5% significance level. If one of these 3 criteria is satisfied, we conclude the draws are convergent.

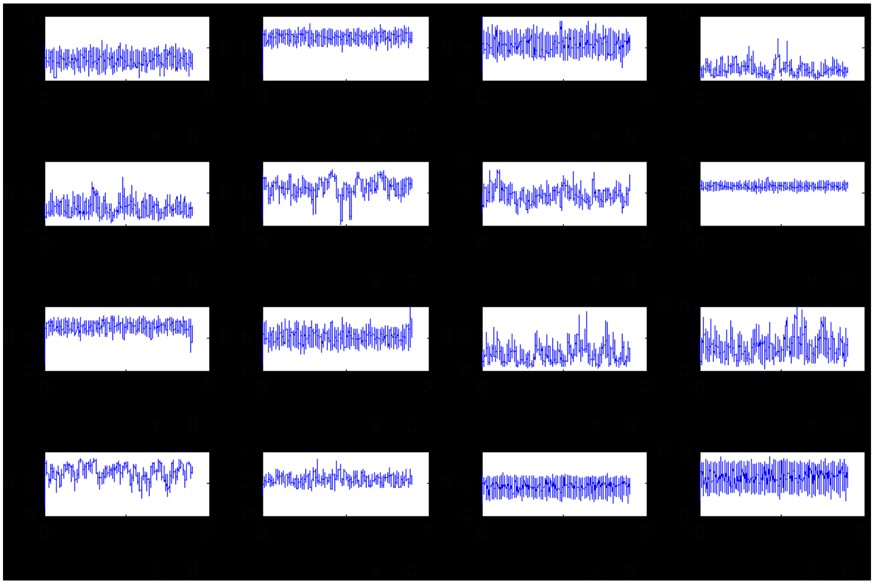

Figure 15: MCMC Draws for Parameters (in the order of Table 14)

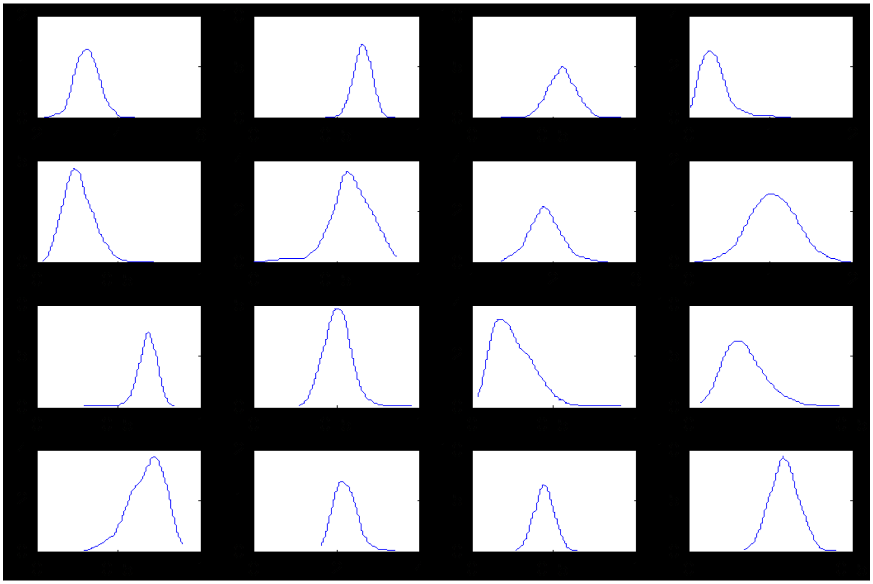

Figure 16: PDFs of Accepted MCMC Draws (in the oder of Table 14)

Parameters	True	Mean	St.Dev.
γ_1	3	3.19	.27
ϕ_1	.7	.66	.05
θ_1	.5	.55	.08
ϖ_1	.3	.36	.14
α_1	.3	.27	.08
β_1	.6	.56	.10
ζ_1	2	1.85	.20
γ_2	.8	1.06	.29
ϕ_2	.65	.67	.06
θ_2	.6	.52	.08
ϖ_2	.2	.24	.11
α_2	.2	.17	.07
β_2	.7	.69	.10
ζ_2	2	2.16	.21
ρ	.5	.44	.06

Table 15: Posterior Summaries

summary statistics are posterior means and standard deviations. We see that the posterior means are close to the true values.

4.5 Empirical application: SP500 and FTSE100

With increasing globalization of the financial markets around the world, we would expect that the returns of the financial instruments in different markets are correlated. If we incorporate such correlation into the prediction of the financial returns in different markets, we may improve the predictability of the returns. Let us take up SP500 and FTSE100 that are two leading stock indicators in U.S. and in Europe, respectively.

As are the cases of many financial return data, the returns on SP500 and on FTSE100 are leptokurtic with robust tails, and thus a bivariate normal distribution may not be appropriate to model the returns.

There have been many empirical studies on the stock indices such as SP500 and FTSE100. For example, Amendola and Stort (2006) studied asymmetries in the conditional mean and variance of the returns on SP500, FTSE100 and NASDAQ, one stock index at a time[32]. Using the EPD distributions, Liu, Wu, and Lee (2004) obtained the maximum likelihood estimates of VaR models of Dow Jones Industrial Average (DJIA), Taiwan Stock Exchange index (TAIEX) and FTSE100, one stock index at a time. Also, there are studies on the correlation structure of stock indices. For example, using Skewed Student-t distribution and quasi-maximum likelihood estimation, Jondeau and Rockinger (2006) estimated a Copula-GARCH model for the four major stock indices in the US, UK, Germany, and France (SP500, FTSE100, DAX and CAC-40). Different from their study, we use a Copula-ARMA-GARCH-EPD model.

We use the monthly data from January 1981 to October 2006 with a total of 319 observations. The data are taken from Global Financial Data. We use these observations to estimate the parameters of Copula-ARMA-GARCH-EPD. The returns are calculated as the monthly rates of change.

The descriptive statistics of these data are shown in Table 16. The means

[32] For more research on SP 500 and/or FTSE100, see Taylor(2004), Bhardwaj and Swanson (2005), Savva et.al. (2005), Bao et. al. (2006), Hu(2006), and Li *et. al.*(2006) .

	FTSE100	SP500
Mean	.0085	.0086
St.Dev.	.0462	.0431
Skewness	−.9050	−.5895
Kurtosis	6.6729	5.4560
Min	.0121	.0102
Max	.1443	.1318
Correlation	.7058	

Table 16: Descriptive Statistics for FTSE100 and SP500

are close to 0. The returns on SP500 and FTSE100 are slightly skewed to the left and they are leptokurtic indicating that the tails of the distributions are fatter than those of the normal distribution. The correlation is about .7058.

The posterior means and standard deviations of the parameters are given in Table 17. We observe that the posterior mean of the correlation parameter is around .6893.

4.6 Chapter Conclusions

A copula is a convenient vehicle to obtain a joint distribution from marginal distributions. In this chapter, we first design a Gaussian copula model with ARMA-GARCH-EPD error terms (i.e., Copula-ARMA-GARCH-EPD). And then, we design Markov Chain Monte Carlo (MCMC) algorithms to estimate this model. Monthly data on SP500 and FTSE100 are analyzed. Results show MCMC algorithm is convergent. The correlation parameter ρ for SP500

Parameters	Mean	St.Dev.
γ_1	$-.0231$	.0000
ϕ_1	$-.9853$	.0170
θ_1	.9893	.0109
ϖ_1	.0024	.0005
α_1	.0527	.0416
β_1	.1734	.1297
ζ_1	2.2455	.1312
γ_2	.0033	.0032
ϕ_2	.9134	.0484
θ_2	$-.9205$	.0533
ϖ_2	.0016	.0003
α_2	.0893	.0635
β_2	.1276	.0938
ζ_2	1.6710	.0675
ρ	.6893	.0422

Table 17: Posterior Means and Standard Deviations of the Parameters

and FTSE100 is .6893.

One of the future extensions is to compare Copula-ARMA-GARCH-EPD model with other nonlinear models. Another extension could add a regime switch parameter in Copula-ARMA-GARCH-EPD model (see Dueker et al. (2007)). Or we can add some macro variables of Bernake et al.(2005)) into Copula-ARMA-GARCH-EPD model and then compare its forecastibility with other models. Last but not the least, one may use Copula-ARMA-GARCH-EPD to price Collateralized Debt Obligations (CDOs).

5 Appendices

Appendix i shows the MCMC algorithm with non-random walk draw for model 1 (MSW-ARMA-GARCH-S_t-in-variance). In Appendix ii, the MCMC algorithm with random walk draw for model 1 is given. Appendix iii shows the MCMC algorithm for Copula-ARMA-GARCH-EPD. Appendix iv presents the simulation results for Copula-EPD and Copula-NORMAL.

5.1 Appendix i: MCMC algorithm using non-random walk draw

Generate S

S is generated by single move method. The distribution of $p(s_t|Y, S_{-t}, \delta)$ can be derived as follows[33]:

$$\begin{aligned} p(s_t|Y, S_{-t}, \delta) &\propto p(s_{t+1}|s_t, \delta_1)p(s_t|s_{t-1}, \delta_1) \\ &\quad * p(y_t|Y_{t-1}, S, \delta)...p(y_T|Y_{T-1}, S, \delta) \\ &\propto p(s_{t+1}|s_t, \delta_1)p(s_t|s_{t-1}, \delta_1) \\ &\quad * \prod_{s=t}^{T}[(\sigma_s^2)^{-1/2}\exp[-\frac{e_s^2}{2\sigma_s^2}]] \end{aligned} \tag{73}$$

After we calculate the $p(s_t = 1|Y, S_{-t}, \delta)$,we compare it with a random number from the uniform distribution. If this probability is greater than the

[33] For more details of derivation for equation (73), please refer to the proof of the Lemma on page 11 of Yoo 2006.

random number , then we set $s_t = 1$, otherwise 0.

Generate $\delta_1 = (p_{00}, p_{11})$

The posterior conditional distribution of p_{00} is :

$$\begin{aligned} p(p_{00}|Y, S, \delta_{-p_{00}}) &\propto p(p_{00}) * p(S|\delta_1) \\ &\propto p_{00}^{u_{00}-1}(1-p_{00})^{u_{01}-1} p_{00}^{n_{00}}(1-p_{00})^{n_{01}} \\ &\propto p_{00}^{u_{00}+n_{00}-1}(1-p_{00})^{u_{01}+n01-1} \end{aligned} \tag{74}$$

That is, we generate p_{00} by Beta distribution,i.e., $p_{00}|S \sim Beta(u_{00} + n_{00}, u_{01} + n_{01})$. Same reasoning, we generate $p_{11}|S$ from Beta$(u_{11} + n_{11}, u_{10} + n_{10})$.

Generate γ

By proper transformation[34], we use following proposal distribution to draw γ:

$$\gamma|Y, X, \Sigma, \delta_{-\gamma} \sim N(\hat{\mu_\gamma}, \hat{\Sigma}_\gamma), \tag{75}$$

Where:

$$\hat{\mu_\gamma} = \hat{\Sigma}_\gamma (X_\gamma^{'} \Sigma^{-1} Y_\gamma + \Sigma_\gamma^{-1} \mu_\gamma), \quad \hat{\Sigma}_\gamma = (X_\gamma^{'} \Sigma^{-1} X_\gamma + \Sigma_\gamma^{-1})^{-1},$$

$$\Sigma = diag\{\sigma_1^2, ..., \sigma_T^2\}, \quad Y_\gamma = [y_1^*, ...y_n^*]', \quad X_\gamma = [x_1^{*'}, ...x_n^{*'}]'.$$

$$y_t^* = y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \theta_i y_{t-i}^*, \quad x_t^* = x_t - \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{i=1}^{q} \theta_i x_{t-i}^*$$

$$e_t \equiv y_t^* - x_t^* \gamma, \quad y_t^* = 0, \quad x_t^* = 0 \quad \text{for all} \quad t \leq 0.$$

[34]See Chib and Greenberg(1994), Yoo (2006), and Das and Yoo (2005).

Generate ϕ

By proper transformation, we draw ϕ using the following proposal distribution:

$$\phi|Y, X, \Sigma, \delta_{-\phi} \sim N(\hat{\mu}_{\phi}, \hat{\Sigma}_{\phi}), \tag{76}$$

Where:

$$\hat{\mu_{\phi}} = \hat{\Sigma}_{\phi}(X_{\phi}^{'}\Sigma^{-1}Y_{\phi} + \Sigma_{\phi}^{-1}\mu_{\phi}), \quad \hat{\Sigma}_{\phi} = (X_{\phi}^{'}\Sigma^{-1}X_{\phi} + \Sigma_{\phi}^{-1})^{-1},$$

$$\Sigma = diag\{\sigma_1^2, ..., \sigma_T^2\}, \quad Y_{\phi} = [\tilde{y}_1, ...\tilde{y}_n]', \quad X_{\phi} = [\tilde{x}_1, ...\tilde{x}_n]'.$$

$$\tilde{y}_t = y_t - x_t\gamma - \sum_{i=1}^{q}\theta_i\tilde{y}_{t-i}, \quad \tilde{x}_t = [\tilde{y}_{t-1}, ..., \tilde{y}_{t-p}]$$

$$e_t = 0, y_t = 0, \tilde{y}_t = 0, \text{for} t \leq 0.$$

Generate θ

We use the following proposal density to draw θ.

$$\theta|Y, X, \Sigma, \delta_{-\theta} \sim N(\hat{\mu}_{\theta}, \hat{\Sigma}_{\theta}), \tag{77}$$

where:

$$\hat{\mu_{\theta}} = \hat{\Sigma}_{\theta}(X_{\theta}^{'}\Sigma^{-1}Y_{\theta} + \Sigma_{\theta}^{-1}\mu_{\theta}), \quad \hat{\Sigma}_{\theta} = (X_{\theta}^{'}\Sigma^{-1}X_{\theta} + \Sigma_{\theta}^{-1})^{-1},$$

$$\Sigma = diag\{\sigma_1^2, ..., \sigma_T^2\}, \quad Y_{\theta} = [\nabla_1\theta^* - e_1(\theta^*), ..., \nabla_n\theta^* - e_n(\theta^*)]^{'},$$

$$X_{\theta} = [\nabla_1^{'}, ..., \nabla_n^{'}]^{'}, \quad \nabla_t = [\nabla_{1t}, ...\nabla_{qt}],$$

$$\nabla_{it} = -e_{t-i}(\theta^*) - \sum_{i=1}^{q}\theta_i^*\nabla_{jt-i}, j = 1, ..., q, \quad \theta^* = arg\min_{\theta}\sum_{t=1}^{n}[e_t(\theta)]^2/\sigma_t^2.$$

Generate α

Same transformation method used for AR coefficient ϕ, we can rewrite (47). The proposal density used to draw α is:

$$\alpha|Y, X, \Sigma, \delta_{-\alpha} \sim N(\hat{\mu}_\alpha, \hat{\Sigma}_\alpha), \tag{78}$$

where:

$$\hat{\mu}_\alpha = \hat{\Sigma}_\alpha(X_\alpha^{'} \wedge^{-1} Y_\alpha + \Sigma_\alpha^{-1}\mu_\alpha), \quad \hat{\Sigma}_\alpha = (X_\alpha^{'} \wedge^{-1} X_\alpha + \Sigma_\alpha^{-1})^{-1},$$

$$\wedge = diag\{2\sigma_1^4, ..., 2\sigma_T^4\}, \quad Y_\alpha = [e_1^2, ..., e_n^2]^{'}, \quad X_\alpha = [\xi_1^{'}, ..., \xi_n^{'}],$$

$$\xi_t = [\eta_t, \tilde{e}_{t-1}^2, ..., \tilde{e}_{t-r}^2], \quad \eta_t = 1 + \sum_{i=1}^{n} \beta_i \eta_{t-i}, \quad \tilde{e}_t^2 = e_t^2 + \sum_{i=1}^{n} \beta_i \tilde{e}_{t-i}^2.$$

Generate β

Similar transformation method as MA coefficient θ, the proposal density to draw β is:

$$\beta|Y, X, \Sigma, \delta_{-\beta} \sim N(\hat{\mu}_\beta, \hat{\Sigma}_\beta), \tag{79}$$

where:

$$\hat{\mu}_\beta = \hat{\Sigma}_\beta(X_\beta^{'} \wedge^{-1} Y_\beta + \Sigma_\beta^{-1}\mu_\beta), \quad \hat{\Sigma}_\beta = (X_\beta \wedge^{-1} X_\beta + \Sigma_\beta^{-1})^{-1},$$

$$\wedge = diag\{2\sigma_1^4, ..., 2\sigma_T^4\}, \quad Y_\beta = [w_1(\beta^*) + \dot{\nabla}_1\beta^*, ..., w_n(\beta^*) + \dot{\nabla}_n\beta^*]^{'},$$

$$X_\beta = [\dot{\nabla}_1^{'}, ..., \dot{\nabla}_n^{'}]', \quad \dot{\nabla}_t = [\dot{\nabla}_{1t}, ..., \dot{\nabla}_{qt}],$$

$$\dot{\nabla}_{it} = -e_{t-i}^2 + w_{t-i}(\beta^*) + \sum_{i=1}^{s} \beta_j^* \dot{\nabla}_{i,t-j}, i = 1, ...s, \beta^* = arg\min_\beta \sum_{t=1}^{T} [w_t(\beta)]^2/2\sigma_t^4.$$

5.2 Appendix ii: MCMC algorithm using random walk draw

Appendix ii gives the proposal density for random walk draw[35]. For S and p_{00}, p_{11}, they are the same as that in Appendix i.

Generate γ

We use following proposal distribution to draw γ:

$$\gamma^{(j)}|Y, X, \Sigma, \delta_{-\gamma} \sim N(\gamma^{(j-1)}, \hat{\Sigma}_{\gamma^{(j-1)}}), \tag{80}$$

where:

$$\hat{\Sigma}_{\gamma^{(j-1)}} = (X'_{\gamma^{(j-1)}} \Sigma^{-1}_{(j-1)} X_{\gamma^{(j-1)}} + \Sigma_{\gamma}^{-1})^{-1},$$

$$\Sigma_{(j-1)} = diag\{\sigma_1^2, ..., \sigma_T^2\},$$

$$Y_\gamma = [y_1^*, ...y_n^*]', \quad X_\gamma = [x_1^{*'}, ...x_n^{*'}]'.$$

$$y_t^* = y_t - \sum_{i=1}^{p} \phi_i y_{t-i} - \sum_{i=1}^{q} \theta_i y_{t-i}^*, \quad x_t^* = x_t - \sum_{i=1}^{p} \phi_i x_{t-i} - \sum_{i=1}^{q} \theta_i x_{t-i}^*$$

$$e_t \equiv y_t^* - x_t^* \gamma, y_t^* = 0, x_t^* = 0 \text{for all} \quad t \leq 0.$$

Generate ϕ

[35] See Qian, Ashizawa and Tsurumi(2004)

We draw ϕ using following proposal distribution:

$$\phi^{(j)}|Y, X, \Sigma, \delta_{-\phi} \sim N(\phi^{(j-1)}, \hat{\Sigma}_{\phi^{(j-1)}}), \quad (81)$$

where:

$$\hat{\Sigma}_{\phi^{(j-1)}} = (X'_{\phi^{(j-1)}}\Sigma^{-1}_{(j-1)}X_{\phi^{(j-1)}} + \Sigma^{-1}_{\phi})^{-1},$$

$$\Sigma_{(j-1)} = diag\{\sigma^2_1, ..., \sigma^2_T\}$$

$$Y_\phi = [\tilde{y}_1, ...\tilde{y}_n]', X_\phi = [\tilde{x}_1, ...\tilde{x}_n]',$$

$$\tilde{y}_t = y_t - x_t\gamma - \sum_{i=1}^{q} \theta_i \tilde{y}_{t-i}, \quad \tilde{x}_t = [\tilde{y}_{t-1}, ..., \tilde{y}_{t-p}]$$

$$e_t = 0, y_t = 0, \tilde{y}_t = 0, \text{for} t \leq 0.$$

Generate θ

The proposal density used to draw θ is:

$$\theta^{(j)}|Y, X, \Sigma, \delta_{-\theta} \sim N(\theta^{(j-1)}, \hat{\Sigma}_{\theta^{(j-1)}}), \quad (82)$$

where:

$$\hat{\Sigma}_{\theta^{(j-1)}} = (X_{\theta^{(j-1)}}\Sigma^{-1}_{(j-1)}X_{\theta^{(j-1)}} + \Sigma^{-1}_{\theta})^{-1},$$

$$\Sigma_{(j-1)} = diag\{\sigma^2_1, ..., \sigma^2_T\},$$

$$X_\theta = [\ddot{x}_1, ..., \ddot{x}_n], \quad \ddot{x} = [\ddot{y}_{t-1}, ..., \ddot{y}_{t-q}],$$

$$\ddot{y}_t = y_t - x_t\gamma + \sum_{i=1}^{p} \phi_i(y_{t-j} - x_{t-j}\gamma) - \sum_{i=1}^{q} \theta_i \ddot{y}_{t-i}$$

$$y_t = 0, \ddot{y}_t = 0, \text{for} t \leq 0.$$

Generate α

The proposal density used to draw α is :

$$\alpha^{(j)}|Y, X, \Sigma, \delta_{-\alpha} \sim N(\alpha^{(j-1)}, \hat{\Sigma}_{\alpha^{(j-1)}}), \tag{83}$$

where:

$$\hat{\Sigma}_{\alpha^{(j-1)}} = (X_{\alpha^{(j-1)}} \wedge_{(j-1)}^{-1} X_{\alpha^{(j-1)}} + \Sigma_{\alpha}^{-1})^{-1},$$

$$\wedge_{(j-1)} = diag\{2\sigma_1^4, ..., 2\sigma_T^4\},$$

$$X_\alpha = [\eta_1, ..., \eta_n], \quad \eta_t = 1 + \sum_{i=1}^{n} \beta_i \eta_{t-i}.$$

Generate β

We use following proposal density to draw β.

$$\beta^{(j)}|Y, X, \Sigma, \delta_{-\beta} \sim N(\beta^{(j-1)}, \hat{\Sigma}_{\beta^{(j-1)}}), \tag{84}$$

where:

$$\hat{\Sigma}_{\beta^{(j-1)}} = (X_{\beta^{(j-1)}} \wedge_{(j-1)}^{-1} X_{\beta^{(j-1)}} + \Sigma_{\beta}^{-1})^{-1},$$

$$\wedge_{(j-1)} = diag\{2\sigma_1^4, ..., 2\sigma_T^4\},$$

$$X_\beta = [x_1^\top, ..., x_n^\top], \quad x_t^\top = -[y_{t-1}^\top, ..., y_{t-s}^\top],$$

$$y_t^\top = e_t^2 - \alpha_0 - \sum_{i=1}^{l} (\alpha_i + \beta_i) e_{t-i}^2 + \sum_{i=1}^{s} \beta_i y_t^\top,$$

$$y_t^\top = 0, \text{for} \quad t \leq 0.$$

5.3 Appendix iii: MCMC algorithm for Copula-ARMA-GARCH-EPD

To draw parameter $\Theta = (\Theta_1, ..., \Theta_M, \rho)$ in Copula-ARMA-GARCH-EPD model, we use following posterior distribution:

$$h(\Theta|Y,\ X) = p(\Theta) * \prod_{i=1}^{n+m} c(F_1(e_{i1})...F_M(e_{iM}); \rho) f_1(\varepsilon_{i1})...f_M(\varepsilon_{iM})$$

Parameters in $\Theta = (\Theta_1, ..., \Theta_M, \rho)$ are $\Theta_i = \{b_i,\ \{\phi_{ij}\}_{j=1...p},\ \{\theta_{ij}\}_{j=1...q},$ $\{\alpha_{ij}\}_{j=0...r}, \{\beta_{ij}\}_{j=1...s}, \zeta_i\}$, $i = 1, ..., M$. We divide parameters $\Theta = (\Theta_1, ..., \Theta_M, \rho)$ into $M + 1$ blocks: Θ_i $(i = 1, ..., M)$ and ρ.

Then, for each parameter Θ_i, the MCMC algorithm is[36]:

1. Draw b_i from independent multivariate Normal by the MH algorithm using the method of random walk draw .

2. Draw each parameter in $\{\phi_{ij}\}_{j=1...p}$ by MH algorithm using random walk draw.

3. Draw each parameter in $\{\theta_{ij}\}_{j=1...q}$ by MH algorithm using random walk draw.

4. Draw each parameter in $\{\alpha_{ij}\}_{j=0...r}$ by MH algorithm using random walk draw.

[36] For more detailed information, one can refer to the algorithm for ARMA-GARCH error term in Nakatsuma(1998) and Qian et al.(2005).

5. Draw each parameter in $\{\beta_{ij}\}_{j=1...s}$ by Efficient Jump algorithm.

6. Draw each parameter in ζ_i by Efficient Jump algorithm.

For parameter ρ, the algorithm is:

Draw covariance matrix $\Sigma = \begin{bmatrix} \sigma_{11}, ..., \sigma_{1M} \\ ... \\ \sigma_{M1}, ..., \sigma_{MM} \end{bmatrix}$ from Inverted Wishart by Gibbs Sampling and calculate $\delta_4 = \rho = D^{-1}\Sigma D^{-1}$, where $D = \begin{bmatrix} \sqrt{\sigma_{11}}, ..., 0 \\ ... \\ 0, ..., \sqrt{\sigma_{MM}} \end{bmatrix}$.

5.4 Appendix iv: Simulation results for Copula-EPD and Copula-NORMAL

Setting $M = 2$ we run following simulation:

$$
\begin{aligned}
&Y_1 = X_1\beta_1 + \varepsilon_1 \qquad (85)\\
&Y_2 = X_2\beta_2 + \varepsilon_2\\
&\varepsilon_1 \sim EPD(\sigma_1, \alpha_1),\ \varepsilon_2 \sim EPD(\sigma_2, \alpha_2).\\
&T = 200\\
&X_i = [1, ..., 1]',\ i = 1, 2.\\
&\beta_1 = .8,\ \beta_2 = .7,\ \sigma_1 = \sigma_2 = 1,\ \alpha_1 = \alpha_2 = 1\\
&\rho = \begin{bmatrix} 1 & .2 \\ .2 & 1 \end{bmatrix}
\end{aligned}
$$

The surface graph of the joint density of Gaussian copula with EPD[37] margins is presented in Figure 17. This density is different from the density of the multivariate EPD distribution in Figure 1 of Gòmez *et.al.* (1998). In Gòmez *et.al.* (1998), the multivariate EPD distribution has only one shape

[37] For density of EPD , please refer to Tsurumi and Shimizu(2006). $EPD(\sigma_i, \alpha_i)$ becomes the normal distribution when $\alpha_i = 2$. If ε_i is distributed as $EPD(\sigma_i, \alpha_i)$, then the PDF of $f_i(\varepsilon_i)$ is:

$$
f_i(\varepsilon_{it}) = \frac{1}{2^{1+1/\alpha_i}\sigma_i\ \Gamma(1+1/\alpha_i)} \exp\left\{-\frac{1}{2}\left|\frac{\varepsilon_{it}}{\sigma_i}\right|^{\alpha_i}\right\} \qquad (86)
$$
$$
0 < \sigma_i < \infty,\ 0 < \alpha_i < \infty,\ i = 1, ..., M.
$$

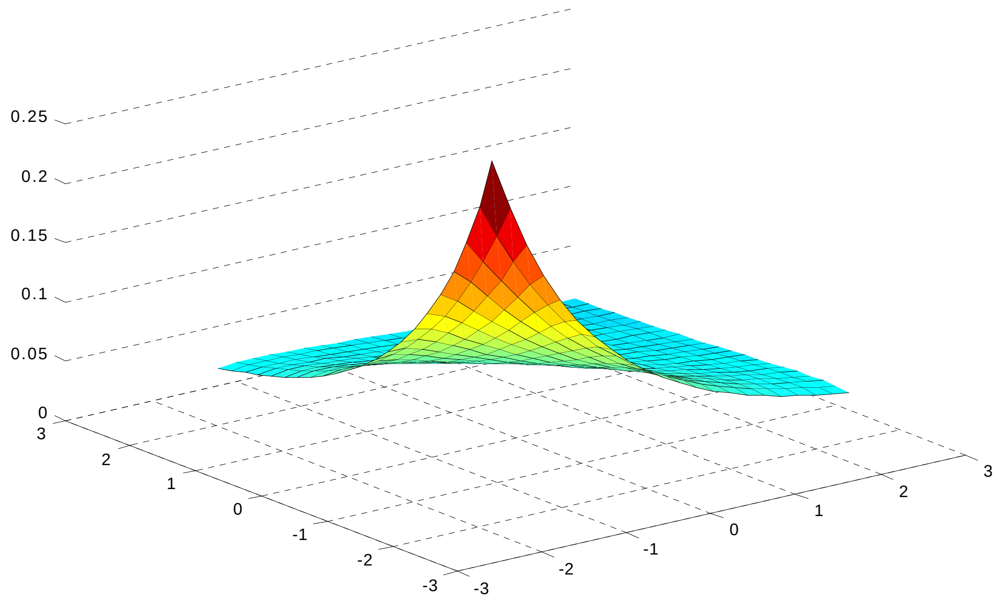

Figure 17: Density of Gaussian Copula with EPD Margins

parameter α , whereas in the Gaussian copula with EPD margins the shape parameter can be different for each marginal distribution.

To verify that the joint density of Gaussian copula with normal margins is multivariate normal, we generated data by setting $\alpha_1 = \alpha_2 = 2$ while keeping the other parameter settings as the same as before. The surface of the joint density of the Gaussian copula with normal margins is presented in Figure 18.

In conclusion, simulation results show that the joint density is far from normal even if we use Gaussian copula with error term ε_i $(i = 1, 2)$ drawn

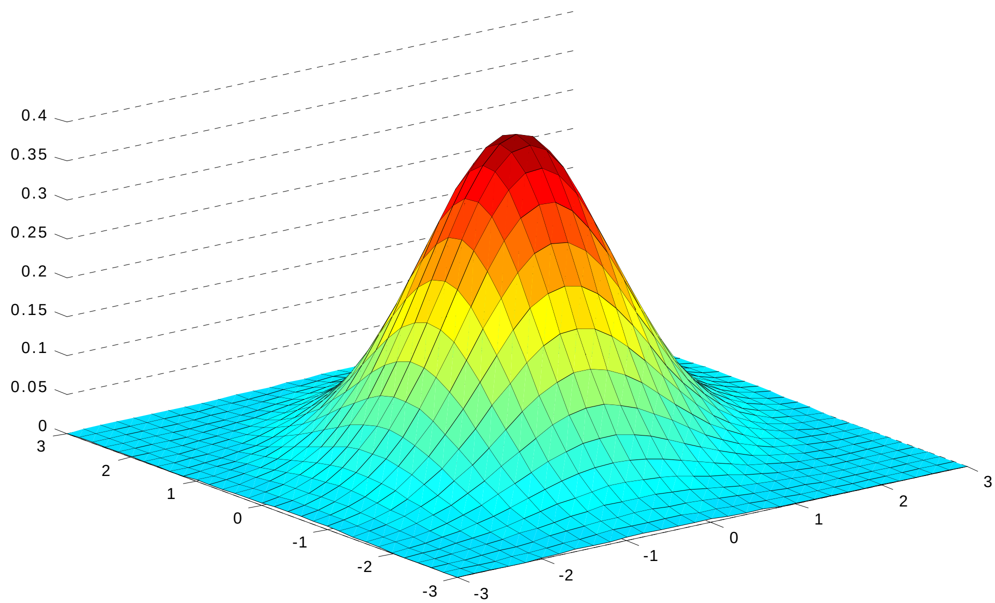

Figure 18: Density of Gaussian Copula with Normal Margins

from exponential power distribution (i.e., Copula-EPD). If the ε_i $(i = 1, 2)$ are drawn from normal distributions (i.e., Copula-NORMAL) then Gaussian copula yields the bivariate normal distribution.

References

AHN, D.-H., AND B. GAO (1999): "A Parametric Nonlinear Model of Term Stucture Dynamics," *The Review of Financial Studies*, 12, 721–762.

AMENDOLA, A., AND G. STORTI (2006): "A Non-linear Time Series Approach to Modelling Asymmetry in Stock Market Indexes," *Statistical Methods and Applications*, 11, 201–216.

ANDERSEN, T., AND T. BOLLERSLEV (1997): "Intraday Periodicity and Volatility Persistence in Financial Markets," *Journal of Empirical Finance*, 4, 115–158.

ANDERSEN, T., AND J. LUND (1997): "Estimating Continuous-time Stochastic Volatility Models of the Short-term Interest Rate," *Journal of Econometrics*, 77, 343–377.

ANTONIOU, A., A. BERNALES, AND D. BEUERMANN (2006): "The Dynamics of the Short-term Interest Rate in the UK," *Journal of Financial Risk Management.*

ARAKELIAN, V., AND P. DELLAPORTAS (2006): "A Threshold Model for Copulas," *Working paper.*

AÏT-SAHALIA, Y. (1996): "Nonparametric Pricing of Interest Rate Derivative Securities," *Econometrica*, 64, 527–560.

——— (1999): "Transition Densities for Interest Rate and other Nonlinear Diffusions," *The Journal of Finance*, 54(4), 1361–1395.

——— (2002): "Maximum Likelihood Estimation of Discretely Sampled Diffusions: a Closed -form Approximation Approach," *Journal of Econometrica*, 70(1), 223–262.

BACON, D., AND D. WATTS (1971): "Estimating the Transition between two Intersecting Straight Lines," *Biometrika*, 58, 525–34.

BALI, T. G. (2000): "Testing the Empirical Performance of Stochastic Volatility Models of the Short-term Interest Rate," *The Journal of Financial and Quantitative Analysis*, 35, 191–215.

BALI, T. G., AND L. WU (2006): "A Comprehensive Analysis of the Short-term Interest-rate Dynamics," *Journal of Banking and Finance*, 30, 1269–1290.

Ball, C. A., and W. N. Torous (1999): "The Stochastic Volatility of Short-term Interest Rates: Some International Evidence," *The Journal of Finance*, LIV(6), 2339–2359.

Bandi, F. M., and T. H. Nguyen (2003): "On the Functional Estimation of Jump-diffusion Models," *Journal of Econometrics*, 116, 293–328.

Barnett, G., R. Kohn, and S. Sheather (1996): "Bayesian Estimation of an Autoregressive Model using Markov Chain Monte Carlo," *Journal of Econometrics*, 74, 237–254.

Barone-Adesi, G., W. Allegretto, E. Dinenis, and G. Sorwar (1999): "Valuation of Derivatives Based on CKLS Interest Rate Models," *http://www.lu.unisi.ch/biblioteca/Pubblicazioni/ArchivioWPeco/wp9906.pdf*.

Bates, D. S. (1996): "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options," *The Review of Financial Studies*, 9, 69–107.

Bernanke, B., J. Boivin, and P. S. Eliasz (2005): "Measuring the Effects of Monetary Policy: A Factor-augmented Vector Autoregressive (FAVAR) Approach," *The Quarterly Journal of Economics*, 120 (1), 387– 422.

Bhardwaj, G., and N. R. Swanson (2005): "An Empirical Investigation of the Usefulness of ARFIMA Models for Predicting Macroeconomic and Financial Time Series," *Journal of Econometrics*, 131 (1-2), 539–578.

Black, F., and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economics*, 81, 637–659.

Blattberg, R. C., and N. J. Gonedes (1974): "A Comparison of the Stable and Student Distributions as Statistical Models for Stock Prices," *Journal of Business*, 47, 244–280.

Blyth, S. (1996): "Out of line," *Risk*, 9 (10), 82–84.

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

——— (1987): "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *The Review of Economics and Statistics*, 69, 542–547.

Boothe, P., and D. Glassman (1985): "The Statistical Distribution of Exchange Rates: Empirical Evidence and Economic Implications," *Research paper, Department of Economics, University of Alberta*, No. 85-22.

Box, G., and G. Tiao (1973): "Bayesian Inference in Statistical Analysis," *Reading Mass: Addison-Wesley.*

Brailsford, T. J., and K. Maheswarah (1998): "The Dynamics of the Australian Short-Term Interest Rate," *SSRN: http://ssrn.com.*

Brennan, M. J., and E. S. Schwartz (1980): "Analyzing Convertible Bonds," *Journal of Financial and Quantitative Analysis*, 15, 907–373.

Brenner, R. J., R. H. Harjes, and K. F. Kroner (1996): "Another Look at Models of the Short-term Interest Rate," *The Journal of Financial and Quantitative Analyst*, 31, 85–107.

Brooks, C. (1996): "Testing for Nonlinearity in Daily Sterling Exchange Rates," *Applied Financial Economics*, 6, 307–17.

Broze, L., O. Scaillet, and J. Zakoian (1995): "Testing for Continuous - time Models of the Short-term Interest Rate," *Journal of Empirical Finance*, 2, 199–223.

Byers, S., and K. Norman (1998): "Forecasting UK and US Interest Rate Using Continuous Time Term Structure Models," *International Review of Financial Analysis*, 7(3), 191–206.

Cai, J. (1994): "A Markov Model of Switching-regime ARCH," *Journal of Business and Economic Statistics*, 12(3), 309–16.

Chan, K., G. Karolyi, F. Longstaff, and A. Sanders (1992): "An Empirical Comparison of Alternative Models of the Short-term Interest Rate," *Journal of Finance*, 47, 1209–1228.

Chan, K. S., and H. Tong (1986): "On Estimating Thresholds in Autogregressive Models," *Journal of Time Series Analysis*, 7(3), 179 – 190.

Chappell, D., J. Padmore, P. Mistry, and C. Ellis (1996): "A Threshhold Model for the French Franc/Deutschmark Exchange Rate," *Journal of Forecasting*, 15, 155–64.

Chen, X., Y. Fan, and A. Patton (2004): "Simple Tests for Models of Dependence between Multiple Financial Time Series, with Applications to U.S. Equity Returns and Exchange Rates," *FMG/IAM,Discussion Paper.*

Chen, X., Y. Fan, and V. Tsyrennikov (2005): "Efficient Estimation of Semiparametric Multivariate Copula Models," *Journal of the American Statistical Association*, forthcoming.

Chernov, M., A. R. Gallan, E. Ghysels, and G. Tauchen (2003): "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225–257.

Christiansen, C. (2008): "Mean Reversion in US and International Short Rates," *Available at SSRN: http://ssrn.com/abstract=1212529.*

Clark, P. (1973): "A Subordinate Stochastic Process Model with Finite Variance for Speculative Prices," *Econometrica*, 41(1), 135–155.

Conley, T., L. Hansen, E. Luttmer, and J. Scheinkman (1997): "Short-term Interest Rates as Subordinated Diffusions," *Reviews of Financial Studies*, 10, 525–577.

Cont, R., and P. Tankov (2002): "Non-parametric Calibration of Jump-diffusion Option Pricing Models," *Working paper.*

Cox, J. (1975): "Notes on Option Pricing I: Constant Elasticity of Variance Diffusion," *Unpublished Note, Standford University, Graduate School of Business.*

Cox, J. C., J. Ingersoll, and S. Ross (1980): "An Analysis of Variable Rate Loan Contracts," *Journal of Finance*, 35, 389–403.

——— (1985): "A Theory of the Term Structure of Interest Rate," *Econometrica*, 53, 373–384.

Crato, N., and P. de Lima (1994): "Long-memory and Nonlinearity: a Time Series Analysis of Stock Returns and Volatilities," *Managerial Finance*, 20, 49–67.

Dahlquist, M. (1996): "On Alternative Interest Rate Processes," *Journal of Banking and Finance*, 20, 1093–1119.

Das, D., and B. H. Yoo (2004): "A Bayesian MCMC Algorithm for Markov Switching GARCH Models," *Unpublished paper.*

Das, S. R. (2002): "The Surprise Element: Jumps in Interest Rates," *Journal of Econometrics*, 106, 27–65.

Demirtas, K. (2006): "Nonlinear Asymmetric Models of the Short-term Interest Rate," *The Journal of Futures Markets*, 26, 869–894.

DOTHAN, U. L. (1978): "On the Term Structure of Interest Rates," *Journal of Financial Economics*, 6, 59–69.

DUEKER, M. J. (1997): "Markov Switching in GARCH Processes and Mean Reverting Stock Market Volatility," *Working paper.*

DUEKER, M. J., M. SOLA, AND F. SPAGNOLO (2007): "Contemporaneous Threshold Autoregressive Models: Estimation, Testing and Forecasting," *Journal of Econometrics*, 141, 517–547.

ELERIAN, O., S. CHIB, AND N. SHEPHARD (2001): "Likelihood Inference for Discretely Observed Nonlinear Diffusions," *Econometrica*, 69, 959–993.

EMBRECHTS, P., A. MCNEIL, AND D. STRAUMANN (1999): "Correlation: Pitfalls and Alternatives," *Working paper.*

ENGLE, R. F. (1982): "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, 50, 987–1008.

EOM, Y. (1998): "On Efficient GMM Estimation of Continuous-time Asset Dynamics: Implications for the Term Structure of Interest Rate," *Working paper, Federal Reserve Bank of New York*.

ERAKER, B. (2001): "MCMC Analysis of Diffusion Models with Application to Finance," *Journal of Business and Economic Statistics*, 19, 177–191.

ERAKER, B., M. JOHANNES, AND N. POLSON (2003): "The Impact of Jumps in Volatility and Returns," *The Journal of Finance*, 58, 1269–1300.

FAFF, R., AND P. GRAY (2006): "On the Estimation and Comparison of Short-Rate Models Using the Generalised Method of Moments," *Journal of Banking and Finance*, 30, 3131–3146.

FAMA, E. (1965): "The Behavior of Stock Market Prices," *Journal of Business*, 38 (1), 34–105.

FAN, J., J. JIANG, AND Z. ZHOU (2003): "Time-dependent Diffusion Models for Term Structure Dynamics," *Statistica Sinica*, 13, 965–992.

FIELITZ, B. (1983): "Stable Distributions and the Mixture of Distributions Hypothesis for Common Stock Returns," *Journal of the American Statistical Association*, 78, 28–36.

Gelfand, A., and D. K. Dey (1994): "Bayesian Model Choice: Asymptotics and Exact Calculations," *Journal of the Royal Statistical Society*, 56, 501–514.

Goldman, E. (2005): "Bayesian Analysis of a Multiple Threshold ARMA Model with CKLS-GARCH Volatility," *International Workshop/Conference on Bayesian Statistics and its Applications.*

Goldman, E. V., and H. Tsurumi (2007): "Kolmogorov-Smirnov, Fluctuation and Z tests for Convergence of Markov Chain Monte Carlo Draws," *Forthcoming in Communications in Statistics, Simulations and Computations.*

Granger, C., and T. Terasvirta (1993): "Modelling Nonlinear Economic Relationships," *Oxford: Oxford University Press.*

Gray, S. F. (1996): "Modeling the Conditional Distribution of Interest Rate as a Regime-switching Process," *Journal of Financial Economics*, 42, 27–62.

Griffin, J., and M. Steel (2006): "Inference with Non-Gaussin Ornstein-Uhlenbeck Processes for Stochastic Volatility," *Journal of Econometrics*, 134, 605–644.

Hamilton, J. D. (1989): "A New Approach to the Economic Analysis of Non-stationary Time Series and the Business Cycle," *Econometrica*, 57(2), 357–84.

Hamilton, J. D., and R. Susmel (1994): "Autoregressive Conditional Heteroskedasticity and Changes in Regime," *Journal of Econometrics*, 64, 307–333.

Henneke, J. S., S. Rachev, and F. J. Fabozzi (2006): "MCMC Based Estimation of Markov Switching ARMA-GARCH Model," *Working paper.*

Hinich, M., and D. Patterson (1985): "Evidence of Nonlinearity in Daily Stock Returns," *Journal of Business and Economic Statistics*, 3, 69–77.

Hiraki, T., and N. Takezawa (1997): "How Sensitive is Short-term Japanese Interest Rate Volatility to the Level of the Interest Rate," *Economics Letters*, 56(3), 325–332.

Hong, S. Y. (2005): "Application of CKLS Model with ARMA-GARCH Errors to Weekly Data on Foreign Exchange Rates," *Master Essay, Rutgers University.*

Hsieh, D. (1989): "Testing for Nonlinear Dependence in Daily Foreign Exchange Rates," *Journal of Business*, 62, 339–69.

——— (1991): "Chaos and Nonlinear Dynamics: Application to Financial Markets," *Journal of Finance*, 46, 1839–77.

JAGANNATHA, R., A. KAPLIN, AND S. SUN (2003): "An Evaluation of Multifactor CIR Models Using LIBOR, Swap Rates, and Cap and Swaption Prices," *Journal of Econometrics*, 116, 113–146.

JIANG, G., AND J. KNIGHT (1997): "A Nonparametric Approach to the Estimation of Diffusion Processes, with an Application to a Short-term Interest Rate Model," *Econometric Theory*, 13, 615–645.

JIANG, G. J. (1998): "Nonparametric Modeling of US Interest Rate Term Stucture Dynamics and Implications on the Prices of Derivative Securities," *The Journal of Financial and Quantitative Analysis*, 33, 465–497.

JONDEAU, E., AND M. ROCKINGER (2006): "The Copula-GARCH Model of Conditional Dependencies: an International Stock Market Application," *Journal of International Money and Finance*, 25 (5), 827–853.

JONES, C. (2003): "Nonlinear Mean Reversion in the Short-term Interest Rate," *Working paper.*

JONES, M. (1987): "Randomly Choosing Parameters from the Stationarity and Invertibility Region of Autoregressive-moving Average Models," *Applied Statistics(Series C)*, 38, 134–138.

KAUFMANN, S., AND S. FRĹŹHWIRTH-SCHNATTER (2000): "Bayesian Analysis of Switching ARCH Models," *Working paper.*

KLAASSEN, F. (2000): "Improving GARCH Volatility Forecasts with Regime Switching GARCH," *Empirical Economics*, 27, 363–94.

KOEDIJK, K., F. NISSEN, P. SCHOTMAN, AND C. WOLFF (1997): "The Dynamics of Short-term Interest Rate Volatility Reconsidered," *European Finance Review*, 1, 105–130.

KRAGER, H., AND P. KUGLER (1993): "Nonlinearities in Foreign Exchange Markets: a Different Perspective," *Journal of International Money and Finance*, 12, 195–208.

LEBARON, B. (1992): "Some Relationships between Volatility and Serial Correlations in Stock Market Returns," *Journal of Business*, 65, 199–219.

LI, D. (2000): "On Default Correlation: a Copula Function Approach," *Journal of Fixed Income*, 9(4), 43–54.

LI, M., M. BOEHNKE, G. R. ABECASIS, AND P. X.-K. SONG (2006): "Quantitative Trait Linkage Analysis Using Gaussian Copulas," *Genetics*, 173 (4), 2317–27.

LIU, M., C. WU, AND H. LEE (2004): "VaR Estimation with Power EWMA Model–Conservativeness, Accuracy and Efficiency," *Working paper.*

LIU, S.-M., AND B. W. BRORSEN (1995): "Maximum Likelihood Estimation of a GARCH-stable Model," *Journal of Applied Econometrics*, 10, 273–285.

LONGSTAFF, F., AND E. SCHWARTZ (1992): "Interest Rate Volatility and the Term Structure : a Two-factor General Equilibrium Model," *Journal of Finance*, 47, 1259–1282.

MAHEU, J., AND T. MCCURDY (2002): "Nonlinear Features of Realized FX Volatility," *Review of Economics and Statistics*, 84, 668–681.

MANDELBROT, B. (1963): "The Variation of Certian Speculative Prices," *Journal of Business*, 36, 394–419.

MARSH, T., AND E. ROSENFELD (1982): "Stochastic Processes for Interest Rates and Equilibrium Bond Prices," *Journal of Finance*, 38, 635–646.

MERTON, R. C. (1973): "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141–183.

MEUCCI, A. (2006): "Beyond Black-Litterman: Views on Non-normal Markets," *Risk*, 87-92.

NAKATSUMA, T. (1998): "A Markov Chain Sampling Algorithm for GARCH Models," *Studies in Nonlinear Dynamics and Econometrics*, 3.

——— (2000): "Bayesian Analysis of ARMA-GARCH Models: A Markov Chain Sampling Approach," *Journal of Econometrics*, 95, 57–69.

NELSON, D. B. (1991): "Conditional Heteroskedasticity in Asset Returns: a New Approach," *Econometrica*, 59, 347–370.

NOWMAN, K. (1997): "Gaussian Estimation of Single-factor Continuous Time Models of the Term Structure of Interest Rate," *The Journal of Finance*, 52, 1695–1706.

——— (2003): "A Note on Gaussian Estimation of the CKLS and CIR Models with Feedback Effects for Japan," *Asia Pacific Financial Markets*, 10, 275–279.

Nowman, K., and B. Yahia (2008): "Euro and FIBOR Interest Rates: a Continuous Time Modelling Analysis," *International Review of Financial Analysis*, In press.

Patton, A. (2006): "Estimation of Multivariate Models for Time Series of Possibly Different Lengths," *Journal of Applied Econometrics*, 21, 147–173.

Pitt, M., D. Chan, and R. Kohn (2006): "Efficient Bayesian Inference for Gaussian Copula Regression Models," *Biometrika*, 93 (3), 537–554.

Ploberger, W., W. Kramer, and K. Kontrus (1989): "A New Test for Structural Stability in the Linear Regression Model," *Journal of Econometrics*, 40, 307–318.

Praetz, P. (1972): "The Distribution of Share Price Changes," *Journal of Business*, 45, 49–55.

Press, S. (1967): "A Compound Events Model of Security Prices," *Journal of Business*, 40, 317–335.

Priestly, M. (1980): "State-dependent Models: a General Approach to Nonlinear Time Series Analysis," *Journal of Time Series Analysis*, 1, 47–71.

——— (1988): "Nonlinear and Nonstationary Time Series Analysis," *London: Academic Press.*

Qian, X., R. Ashizawa, and H. Tsurumi (2005): "Bayesian Estimation of a Spot Rate Model," *Working paper, Rutgers University.*

——— (2008): "Bayesian, MLE, GMM Estimation of a Spot Rate Model," *Communications in Statistics*, 34, 221–246.

Reno, R. (2005): "Nonparametric Estimation of Stochstic Volatility Models," *Economics Letters*, 90(3), 390–395.

Rodrigues, P. M., and A. Rubia (2005): "The Performance of Unit Root Test under Level-dependent Heteroskedasticity," *Economics Letters*, 89, 262–268.

Romeo, J., T. NI, and P. de Lima AC (2006): "Bivariate Survival Modeling: a Bayesian Approach Based on Copulas," *Lifetime Data Analysis*, 12, 205–222.

Saltoglu, B. (2003): "Comparing Forecasting Ability of Parametric and Nonparametric Methods: Application with Canadian Monthly Interest Rates," *Applied Financial Economics*, 13, 169–176.

SANFORD, A. D., AND G. M. MARTIN (2006): "Bayesian Comparison of Several Continuous Time Models of the Australian Short Rate," *Accounting and Finance*, 46(2), 309–326.

SCHEINKMAN, J. A., AND B. LE BARON (1989): "Nonlinear Dynamics and Stock Returns," *Journal of Business*, 62, 311–37.

SHAW, J. (1997): "Beyond VAR and Stress Testing in VAR: Understanding and Applying Value at Risk," *Risk Publications, London*, pp. 211–224.

SINGLETON, K. J. (2001): "Estimation of Affine Asset Pricing Models Using the Empirical Characteristic Function," *Journal of Econometrics*, 102, 111–141.

SKLAR, A. (1959): "Fonctions de rĺępartition Íd' n dimensions et leurs marges," *Publications de l'Institut de Statistique de L'Universitĺę de Paris*, 8, 229–23.

SMIRNOV, N. (1939): "On the Estimation of the Discrepancy between Empirical Curves of Distribution for two Independent Samples," *Bulletin Mathematique de VUniversite de Moscou*, 2, fsc.2.

SORWAR, G. (2005): "Implied Derivative Security Prices Based on two-factor Interest Model: a UK Application," *Applied Financial Economics*, 15(10), 739–744.

SORWAR, G., G. BARONE-ADESI, AND W. ALLEGRETTO (2007): "Valuation of Derivatives Based on Single-factor Interest Rate Models," *Global Finance Journal*, 18, 251–269.

SPIEGELHALTER, D. J., N. G. BEST, B. P. CARLIN, AND A. VAN DER LINDE (2002): "Bayesian Measures of Model Complexity and Fit (with Discussion)," *Journal of the Royal Statistical Society:Series B*, 64(4), 583–639.

STANTON, R. (1997): "A Nonparametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk," *Journal of Finance*, 52, 1973–2002.

SUBBOTIN, M. (1923): "The Law of Frequency of Error," *Mathematicheskii Sbornik*, 31, 296–301.

TERASVIRTA, T. (1994): "Specification, Estimation and Evaluation of Smooth Transtion Autoregressive Models," *Journal of the American Statistical Association*, 89, 208–18.

——— (1998): "Modelling Economic Relationships with Smooth Transition Regressions.," *Handbook of Applied Economic Statistics. New York: Marcek Dekker*, pp. 507–52.

TONG, H. (1978): "On a Threshold Model," *in C.H. Chen (ed.). Pattern Recognition and Sigmal Processing, Amsterdam: Sijthoff and Noordhoff*, pp. 101–41.

——— (1990): "Non-linear Time Series: A Dynamical Systems Approach," *Oxford:Oxford University Press.*

TONG, H., AND K. LIM (1980): "Threshold Autoregressions, Limit Cycles, and Data," *Journal of the Royal Statistical Society B*, 42, 245–92.

TREEPONGKARUNA, S., AND S. GRAY (2003): "Short-term Interest Rate Models:Valuing Interest Rate Derivatives Using a Monte-Carlo Approach," *Accounting and Finance*, 43(2), 231–259.

TSE, Y. (1995): "Some International Evidence on the Stochastic Behavior of Interest Rates," *Journal of International Money and Finance*, 14(5), 721–738.

TSURUMI, H., AND H. SHIMIZU (2008): "Bayesian Inference of the Regression Model with Error Terms Following the Exponential Power Distribtuion and Unbiasedness of the LAD Estimator," *Communications in Statistics*, 37, 221–246.

VASICEK, O. (1977): "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, 5, 177–188.

WESTERFIELD, R. (1977): "The Distribution of Common Stock Price Changes: An Application of Transactions Time and Subordinate Stochastic Models," *Journal of Financial and Quantitative Analysis*, 12, 743–765.

XU, Y. (2004): "Incoporating Estimation Risk in Copula-based Portfolio Optimization," *Working paper.*

YOO, B. H. (2006): "A Bayesian Analysis of Markov Switching Models with ARMA-GARCH Error," *Working paper, Rutgers University.*

ZIMMER, D., AND P. TRIVEDI (2006): "Using Trivariate Copulas to Model Sample Selection and Treatment Effects: Application to Family Health Care Demand," *Journal of Business and Economic Statistics*, 24(1), 63–76.

Curriculum Vitae

Liuling Li

2009	PhD, Economics, Rutgers University
2008	MS, Math Finance, Rutgers University
2001	MM, Management Science/Finance, Shanghai Jiao Tong University
1998	BE, International Finance, Shanghai Jiao Tong University