The Academy of Economic Studies
The Faculty of Finance, Insurance, Banking and Stock Exchange
Doctoral School of Finance and Banking

Bayesian Analysis of Stochastic Volatility Model

MSc Student: Alexandru Burducea
Supervisor: Professor Moisă Altăr

Bucharest, June 2012

Contents

Motivation and Objectives

- Substantial empirical evidence indicates that most returns of financial assets are not normally distributed. They dispay some properties such as time-varying volatility and volatility clustering.
- The financial time series data such as stock returns and foreign exchange returns have several properties which depart from a normality assumption. Major characteristics of return distributions for financial variables are the skewness, heavy-tailness and volatility clustering with leverage effects.
- These properties are important not only for describing return distributions but also for the asset allocation, option pricing, forecasting and risk management.

- The objective of this dissertation paper is to compare different estimation methods for the classic stochastic volatility model proposed by Taylor (1982).

Literature Review

- Commonly, the errors are assumed Normally distributed. However, the assumption of Normality has been questioned in the literature and heavier tailed distributions have been proposed as alternatives.
- The discrete mixtures of normal distributions (MN) seems to be a very appealing candidate due to its flexibility to model any shape of continuous distribution. Kon (1984) discusses its applications to 30 stocks in Dow-Jones Industrial Average and concludes that MN model has substantially more descriptive validity than the student-t or normal models.
- Kim, Shephard, and Chib (1998) retain the convenient linear form of the state space model and use a prespecified mixture of seven normals.
- Nakajima and Omori (2009) and Jacquier et al. (2004) consider the student distribution, Maheu and Schotman (1998), Maheu and Jensen (2012), Delatola and Griffin (2011) use a mixture of Normals and Abanto-Valle et al. (2010) apply scale mixture of Normals using diferent mixing parameters.

Stochastic Volatility Model

- Returns have a volatility (standard deviation) that changes over time. The main purpose of stochastic volatility models is to describe this underlying time-varying volatility.
- In the standard discrete SV model the volatility follows a latent AR(1) process.
- The first SV model was proposed by Taylor (1982) and has the following form:

$$y_t = e^{h_t/2}\varepsilon_t \qquad \varepsilon_t \sim N(0,1)$$

$$h_t = \omega + \Phi * h_{t-1} + \sigma_w * w_t \qquad w_t \sim N(0,1)$$

- Equivalently

$$\log y_t^2 = h_t + \log \varepsilon_t^2 \qquad \log \varepsilon_t^2 \sim \log \chi_1^2$$

$$h_t = \omega + \Phi * h_{t-1} + w_t \qquad w_t \sim N(0, \sigma^2_w)$$

Stochastic Volatility Model

$y_t{}^* = h_t + z_t$ $\qquad z_t \sim \log\chi_1{}^2$

$h_t = \omega + \Phi * h_{t-1} + w_t$ $\qquad w_t \sim N(0, \sigma^2{}_w)$

- The general form for a Dynamic Linear Model is:

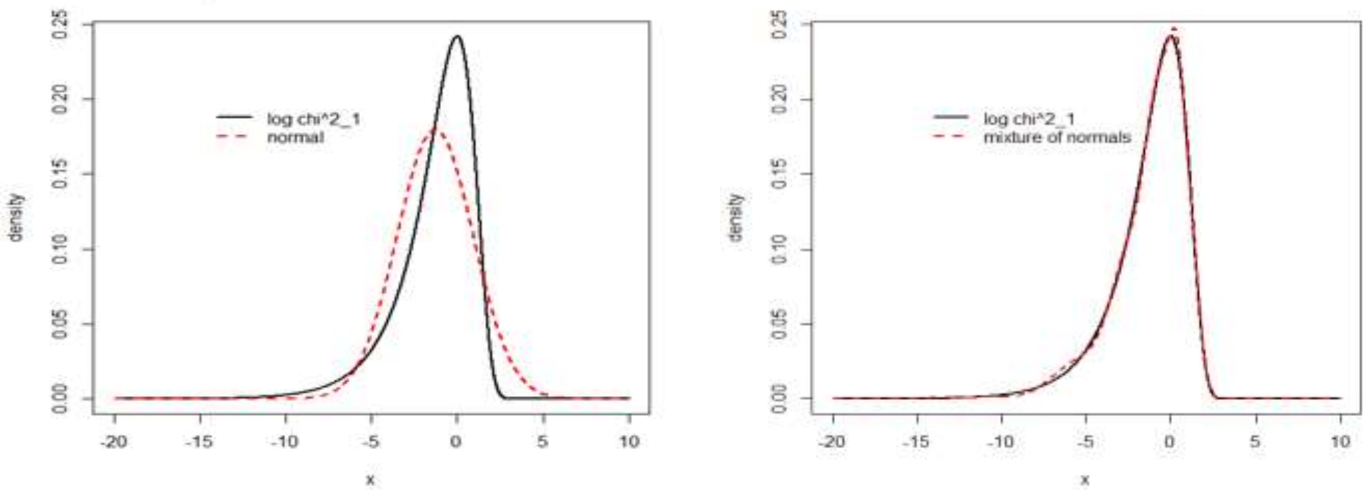
$y_t = F'_t * \beta_t + v_t$ where $\varepsilon_t \sim N(0, V_t)$

$\beta_t = G_t{}^* \beta_{t-1} + w_t$ where $w_t \sim N(0, W_t)$

y_t : sequence of observations
F'_t: vector of explanatory variables
β_t: state vector
G_t: evolution matrix
$\beta_1 \sim N(a_1, R_1)$
v_t independent of w_t

Stochastic Volatility Model

- Even though we can write the SV model as a dynamic linear model $\log \varepsilon_t{}^2$ is non – Gaussian, actually is log chi-squared.
- $E(\log\chi_1{}^2) = -1.27$; $Var(\log\chi_1{}^2) = -4.9$
- The following approximations can be used to estimate the model:



$\log \chi_1^2$ vs $N(-1.27, 4.9)$ and mixture of 7 normals

- A normal distribution is not good enough to approximate the log chi-squared, the mixture of 7 distributions is a better one

Stochastic Volatility Model

$$\log \chi_1^2 \approx \sum_{k=1}^{7} \pi_k N(\mu_k, \tau_k^2)$$

μ_k	-11.400	-9.837	-5.243	-2.358	-0.650	0.524	1.507
τ_k^2	5.795	5.179	2.613	1.262	0.640	0.340	0.167
π_k	0.00730	0.00002	0.10556	0.25750	0.34001	0.24566	0.04395

- Kim, Shephard, and Chib (1998) use a mixture of seven normal distributions instead of a normal distribution.

Stochastic Volatility Model

- The Dirichlet process mixture of normals model (DPM) was introduced by Lo (1984) and Ferguson (1983).

$$f_k(z) = \sum_{j=1}^{k} p_j \mathrm{N}\left(z \left| \mu_j', \sigma_j'^2 \right.\right)$$

Where p_1, p_2,…, p_k follow a Dirichlet distribution. This is a k-component normal mixture model where the location parameter and the scale parameter differ from component-to-component and p_j are the mixing weights. The DPM belongs in the class of infinite mixture models.

- The DPM builds on the well known property that a flexible distribution can be found by mixing together a number of known distributions. It extends this concept by mixing together an infinite number of distributions.

Stochastic Volatility Model

- The DPM possesses a number of attractive features as a Bayesian nonparametric estimator:
  1) as a prior to a infinite order mixture model the DPM is more flexible and realistic than a parametric mixture model with a predetermined number of components,
  2) as a conjugate prior the DPM is easy to use and facilitates Gibbs sampling,
  3) with the DPM model the data determines the correct number of mixture clusters,
  4) we can impose parsimony through the prior and

$$G \sim DP(a, G_0)$$

G_0 is a base distribution and a is a positive scaling parameter

Model Comparison

- The fit of the models will be assessed using log predictive scores as proposed by Kim et al. (1998) and Delatola and Griffin(2011).
- The average log predictive score for one-step ahead prediction is given by

$$\text{LPS} = -\frac{1}{T}\sum_{j=1}^{T} log p(y_j \mid y_{j-1,} \Phi)$$

- The predictive likelihood records the predictive record of a model on a set of data and provides a natural quantity for model comparison (Geweke & Whiteman (2006)).
- Smaller values of the LPS indicate a better fitting model.

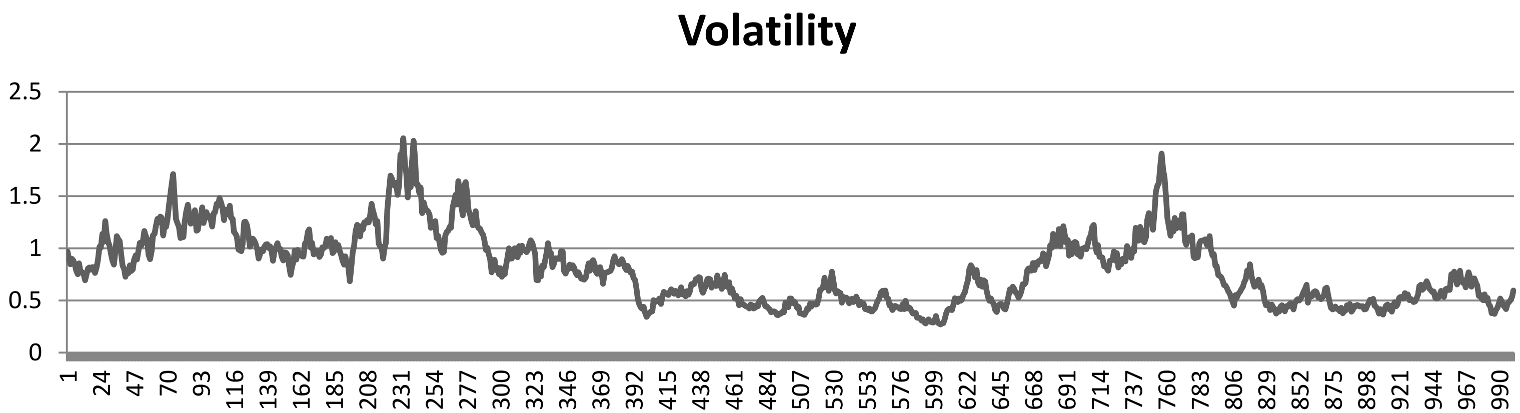
MCMC Algorithm

- Gibbs sampler

- initialize Φ, σ_w , σ_ε,μ_0, μ_i, s
- sample h| y, Φ, σ_w , σ_ε,μ_0, μ_i, s using FFBS
- sample s| y, Φ, σ_w , σ_ε,μ_0, μ_i, h
- sample σ_ε,μ_0, μ_i | y, Φ, σ_w, s ,h
- sample Φ, σ_w | y, σ_ε, μ_0, μ_i,s, h

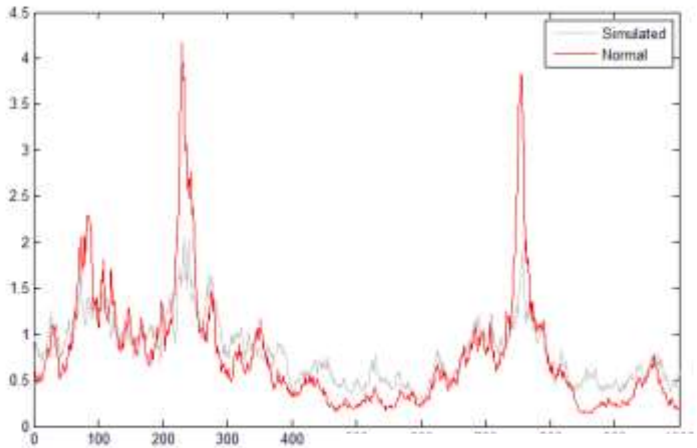
- Conditional on s, the model for y is a Gaussian dynamic linear model and so the log-volatilities h can be updated simultaneously using the forward filtering backward sampling (FFBS) algorithm (Carter and Kohn 1994; Fruhwirth-Schnatter 1994; Durbin and Koopman 2002).

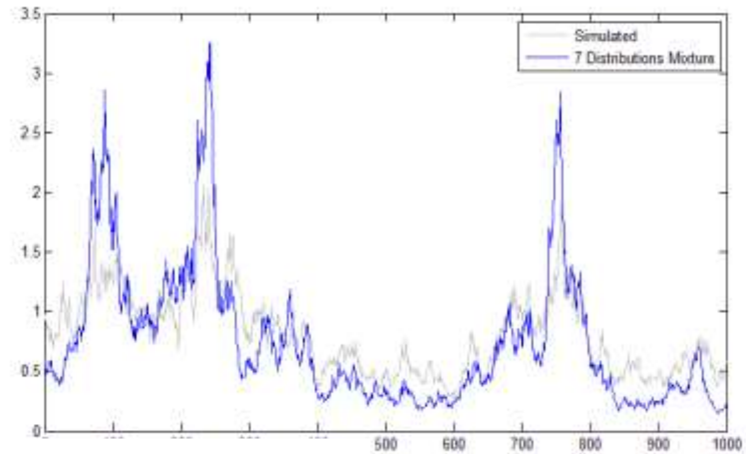
Simulated data

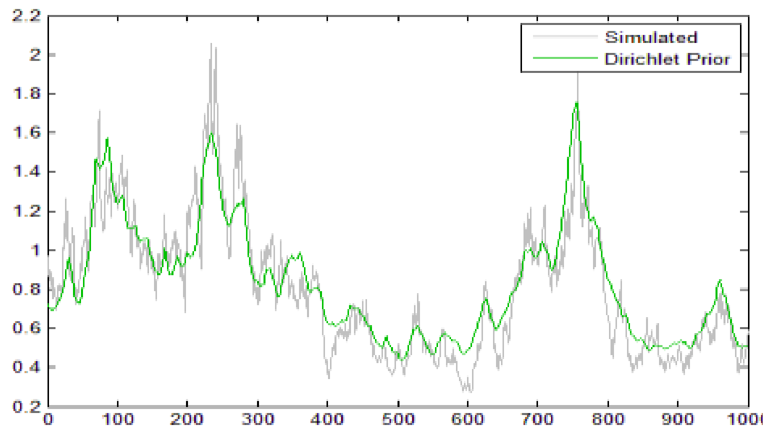


1000 observations, omega=-0.00645, phi=0.15

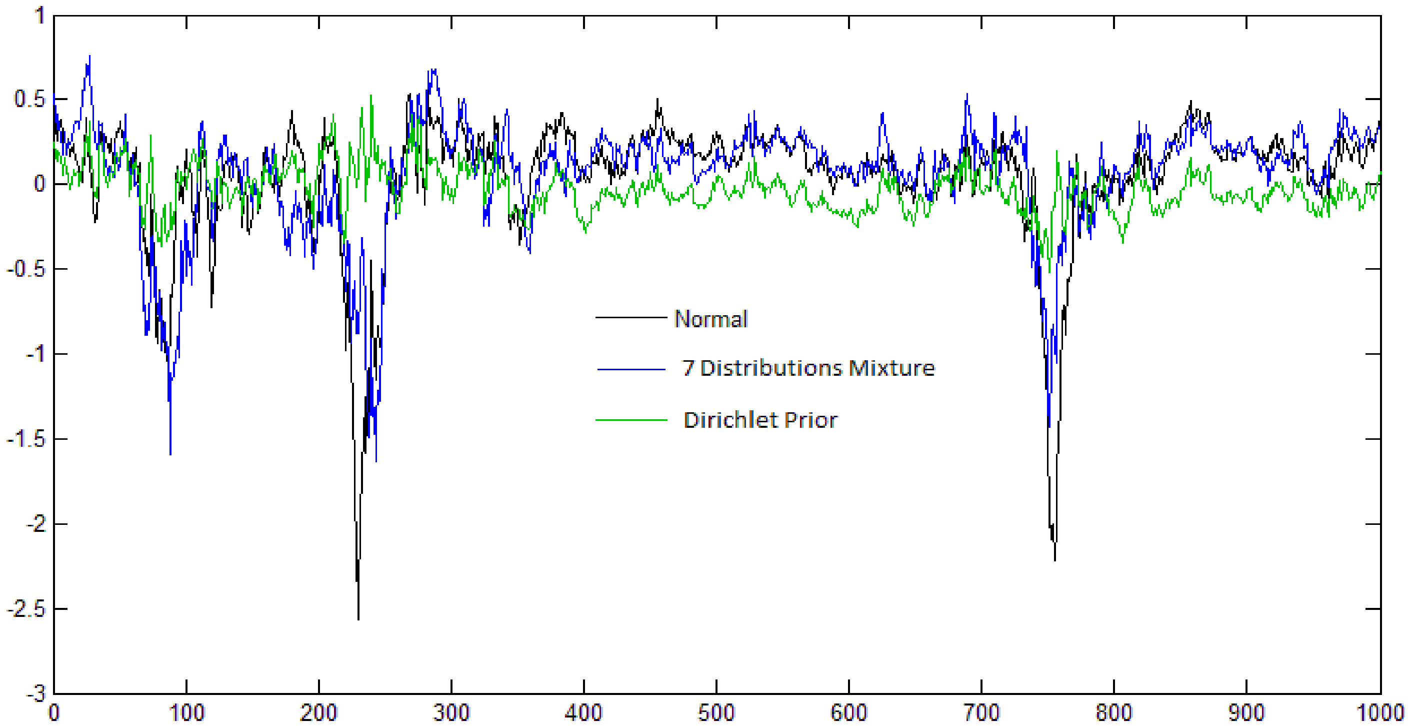
Simulated data







Simulated data



The difference between the simulated data and the result from each estimation method

Simulated data – parameter estimation

Normal Approximation	mean	std	2.5%	97.5%	TRUE
omega	-0.01215	0.007638	-0.0272	0.0026	-0.0065
phi	0.977112	0.005944	0.9651	0.9883	0.9900
sigmaw	0.162134	0.004365	0.1514	0.1713	0.1500

7 Distributions Mixture	mean	std	2.5%	97.5%	TRUE
omega	-0.00943	0.006508	-0.0222	0.0034	-0.0065
phi	0.982153	0.004763	0.9727	0.9914	0.9900
sigmaw	0.134387	0.003666	0.1267	0.1412	0.1500

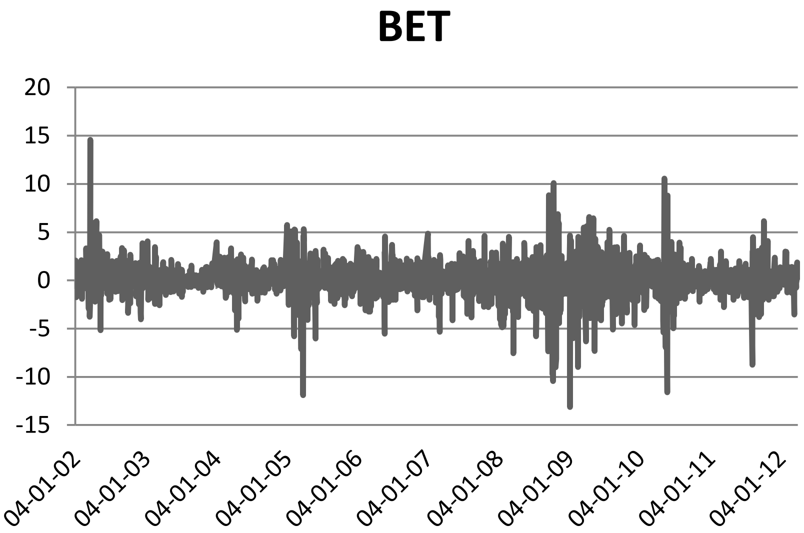
Dirichlet Prior*	mean	std	2.5%	97.5%	TRUE
omega	-0.0054	0.2015	-0.0096	-0.0017	-0.0065
phi	0.9812	0.0091	0.9593	0.9948	0.9900
sigmaw	0.1497	0.009	0.1039	0.2177	0.1500

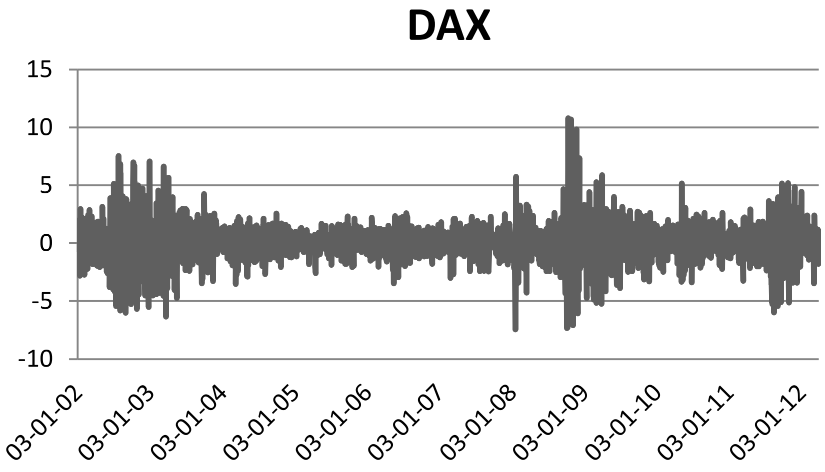
Simulated data	LPS
Normal Approximation	8.9455
7 Distributions Mixture	8.6667
Dirichlet Prior	7.5631

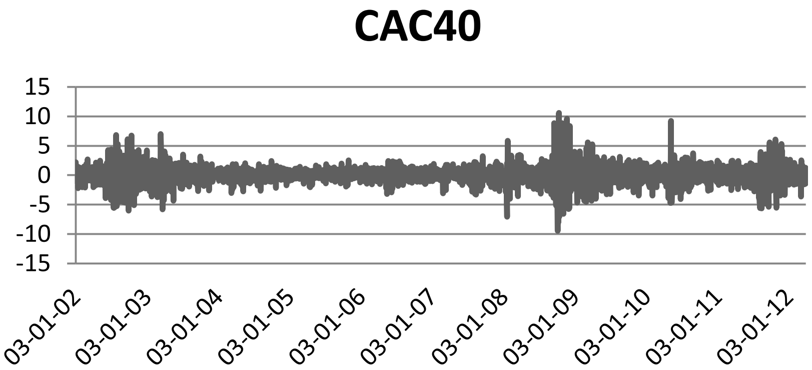
* 27 mixtures (7 - lower 2.5% bound, 65 – upper 97.5% bound)

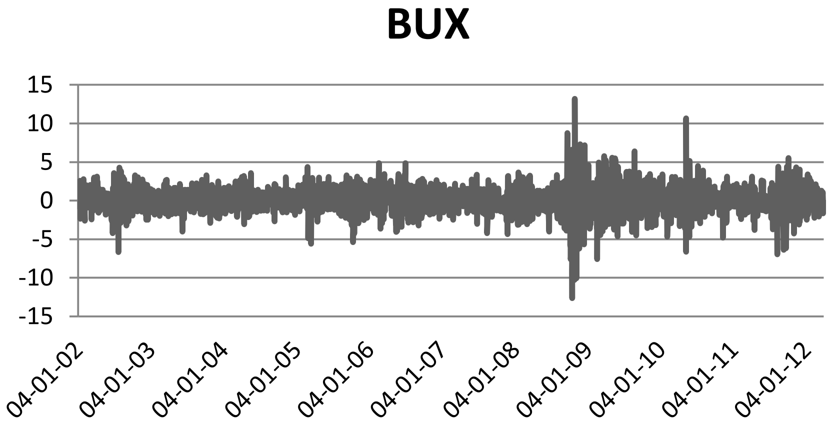
burn in =2,000 obs, iterations = 10,000 obs

Empirical study

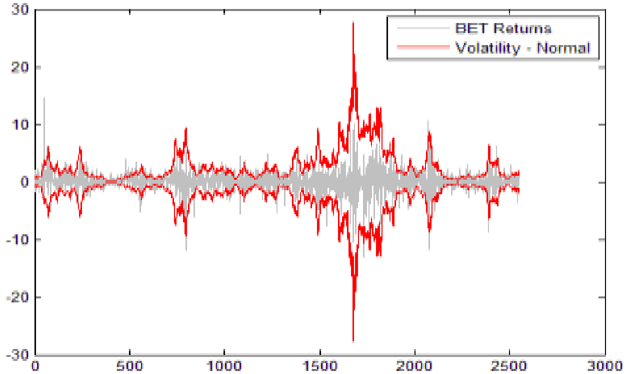


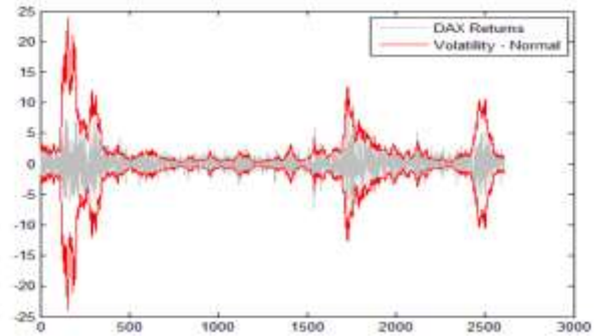


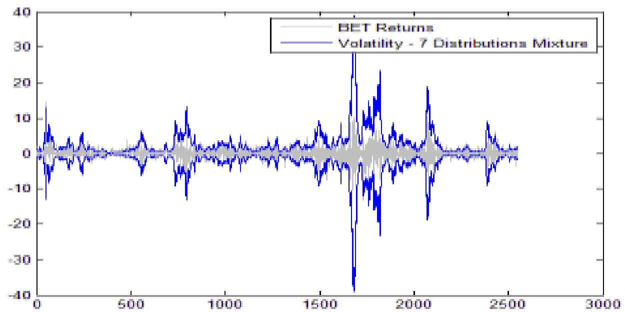


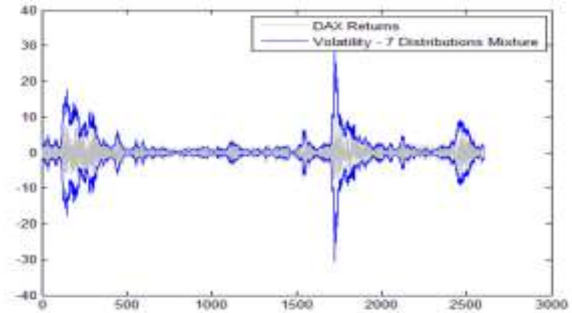


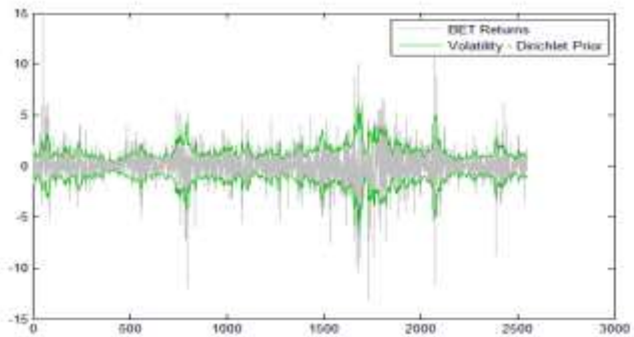
Volatility Estimation

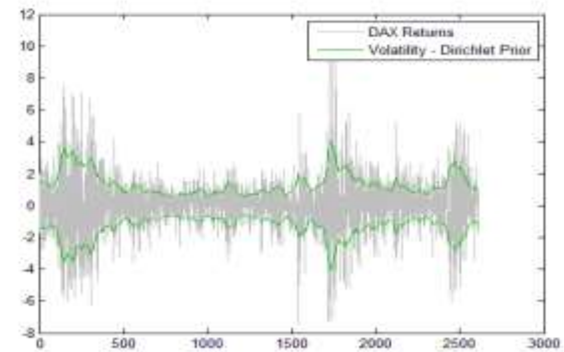




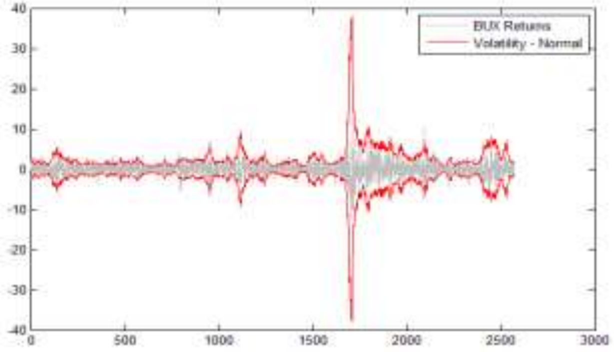


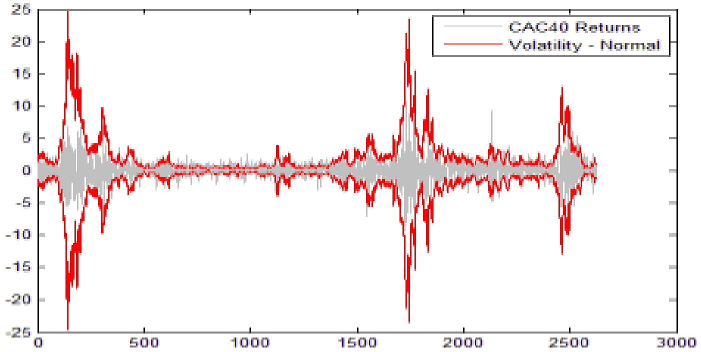


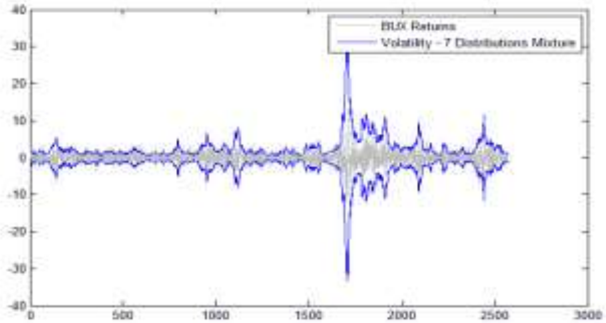


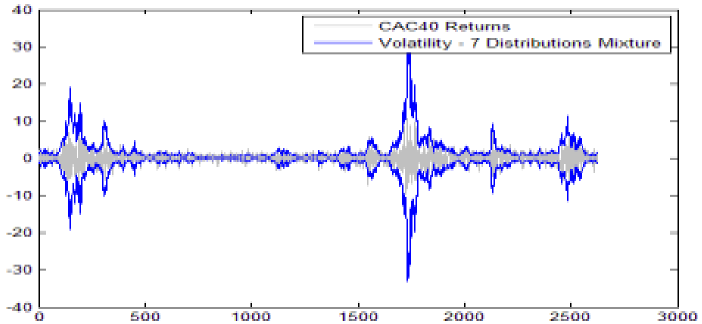


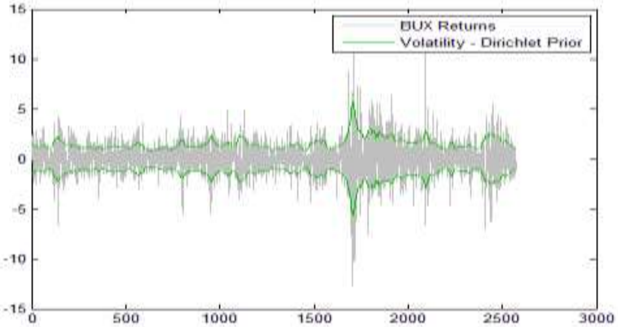
Volatility Estimation

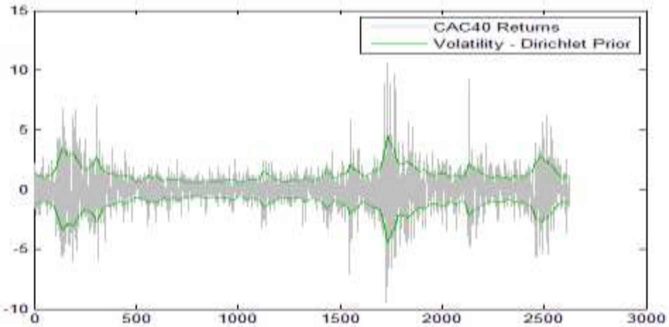












Parameter Estimation

BET

Normal Approximation	mean	std	2.5%	97.5%
omega	0.0065	0.0039	-0.0011	0.0141
phi	0.9875	0.0028	0.9821	0.9929
sigmaw	0.157	0.0022	0.1528	0.1613
7 Distributions Mixture	mean	std	2.5%	97.5%
omega	0.0136	0.0054	0.0028	0.024
phi	0.9764	0.0038	0.9691	0.9836
sigmaw	0.2115	0.0032	0.2052	0.2179
Dirichlet Prior*	mean	std	2.5%	97.5%
omega	0.0058	0.119	0.0035	0.0079
phi	0.946	0.0125	0.9185	0.9683
sigmaw	0.3515	0.0292	0.2685	0.4430

* 10 mixtures (4 - lower 2.5% bound, 23 – upper 97.5% bound)

DAX

Normal Approximation	mean	std	2.5%	97.5%
omega	0.005	0.0033	-0.0017	0.0116
phi	0.9894	0.0027	0.9842	0.9946
sigmaw	0.1423	0.0023	0.1382	0.1472
7 Distributions Mixture	mean	std	2.5%	97.5%
omega	0.0071	0.0039	-0.00057	0.015
phi	0.9853	0.0031	0.979	0.9913
sigmaw	0.1641	0.0067	0.1527	0.1733
Dirichlet Prior*	mean	std	2.5%	97.5%
omega	0.0013	0.467	-0.0082	0.0079
phi	0.9922	0.0033	0.9851	0.998
sigmaw	0.1286	0.0036	0.1026	0.1585

* 12 mixtures (6 - lower 2.5% bound, 26 – upper 97.5% bound)

burn in =2,000 obs, iterations = 10,000 obs

Parameter Estimation

BUX

Normal Approximation	mean	std	2.5%	97.5%
omega	0.017	0.0055	0.0061	0.0277
phi	0.9772	0.0033	0.9709	0.9837
sigmaw	0.1574	0.0029	0.1514	0.1626
7 Distributions Mixture	mean	std	2.5%	97.5%
omega	0.0163	0.0056	0.00570	0.0275
phi	0.9753	0.004	0.9675	0.983
sigmaw	0.1723	0.0073	0.162	0.1893
Dirichlet Prior*	mean	std	2.5%	97.5%
omega	0.00666	0.1443	0.0040	0.0091
phi	0.9797	0.006	0.9668	0.9902
sigmaw	0.1509	0.0059	0.1177	0.1902

* 38 mixtures (6 - lower 2.5% bound, 123– upper 97.5% bound)

CAC40

Normal Approximation	mean	std	2.5%	97.5%
omega	0.0043	0.0033	-0.0023	0.0108
phi	0.9891	0.0027	0.9838	0.9944
sigmaw	0.148	0.002	0.144	0.1520
7 Distributions Mixture	mean	std	2.5%	97.5%
omega	0.0046	0.0036	-0.00240	0.0116
phi	0.9854	0.0032	0.9791	0.9915
sigmaw	0.1645	0.0023	0.1601	0.1690
Dirichlet Prior	mean	std	2.5%	97.5%
omega	0.00417	0.3227	-0.0071555	0.0062765
phi	0.9918	0.0033	0.9844	0.9975
sigmaw	0.1284	0.0036	1.1051	0.1576

* 9 mixtures (5 - lower 2.5% bound, 15– upper 97.5% bound)

burn in =2,000 obs, iterations = 10,000 obs

Parameter Estimation

BET	LPS
Normal Approximation	7.6301
7 Distributions Mixture	6.8265
Dirichlet Prior	6.8095

BUX	LPS
Normal Approximation	7.6982
7 Distributions Mixture	7.1879
Dirichlet Prior	6.9563

DAX	LPS
Normal Approximation	7.9019
7 Distributions Mixture	7.5231
Dirichlet Prior	6.7484

CAC40	LPS
Normal Approximation	7.8574
7 Distributions Mixture	7.2933
Dirichlet Prior	6.805

burn in =2,000 obs, iterations = 10,000 obs

Conclusions

- This paper presents a method for Bayesian inference for the stochastic volatility model. It is modelled log ${\varepsilon_t}^2$ rather than ε_t, as considered by Jensen and Maheu (2010). This allows efficient computational techniques using forward filtering backward sampling methods to be applied to the difficult problem of updating the log volatilities in a Gibbs sampler.
- The results obtained suggest that the semiparametric model has a better output than the model with the normal approximation or the one with a mixture of seven normal distributions. Making certain restrictions on the parameters of the Dirichlet process more parsimony models can be obtained, like the one with normal iid errors or the Student-t SV model of Harvey et al.(1994). An explanation why the Dirichlet SV model proved to better fit the data is because the other models are nested within this one.

Conclusions

- The sampler is slower when a Dirichlet prior is used and the draws from the posterior distributions are more correlated (the correlation coefficient is usually in the interval (- 0.4 ,+ 0.4)). This indicates a slow mixing chain which will need to be run for a longer period in order to obtain good estimates because it takes more time to explore the entire distribution.
- Future work should consider the addition of a leverage effect and further nonparametric modelling of the volatility equation to better fit the characteristics of the financial time series.

References

> Abanto-Valle, C., D. Bandyopadhyay, V. Lachos, and I. Enriquez (2010), “Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions”, Computational Statistics and Data Analysis, 54,pp. 2883-2898.

> Broto, C. and E. Ruiz (2002). “Estimation methods for stochastic volatility models: a survey”, Universidad Carlos III de Madrid, Working Paper 02-54 (14), Statistics and Econometrics Series, November 2002

> Carter, C. K. and R. Kohn (1996), “Markov chain Monte Carlo in conditionally Gaussian state space models”, Biometrika, 83, 3,pp. 589-601

> Chib, S., F. Nardari, and N. Shephard (2002), “Markov chain Monte Carlo methods for stochastic volatility models”, Journal of Econometrics vol. 108, pp.281-316

> Delatola, E.I., and J. E. Griffin (2011), “Bayesian Nonparametric Modelling of the Return Distribution with Stochastic Volatility”, Bayesian Analysis 4, p. 901-926

> Durham, G. B. (2007) , “SV mixture models with application to S&P 500 index return”, Journal of Financial Economics 85, 822–856.

> Escobar, M. D. and M. West (1995), “Bayesian Density Estimation and Inference Using Mixtures”, Journal of the American Statistical Association, 90, pp.577-588

> Fruhwirth-Schnatter, S. (1992), “Data Augmentation and Dynamic Linear Models“, Working Paper, Forschungsberichte, Department of Statistics and Mathematics, WU Vienna University of Economics and Business, Vienna

> Griffin, J. E. (2011), “Default priors for density estimation with mixture models”, Bayesian Analysis 5, 45–64.

> Jacquier, E., N. G. Polson, and P. E. Rossi(1994), “Bayesian Analysis of Stochastic Volatility Models”, Journal of Business & Economic Statistics, 12, pp. 371-389.

> Jensen, M. J. and J. M. Maheu (2010), “Bayesian semiparametric stochastic volatility modeling” Journal of Econometrics, 157, p. 306-316

(2012), “Estimating a Semiparametric Asymmetric Stochastic Volatility Model with a Dirichlet Process Mixture” ,Federal Reserve Bank of Atlanta, Working Paper Series 2012-6

References

> Jong, P., and N. Shepard (1995), “The simulation smoother for time series models”, Biometrika, 82, 2, pp. 339-50

> Kim S., N. Shepard, and S. Chib (1998), “Stochastic Volatility Likelihood Inference and Comparison with ARCH Models”, The Review of Economic Studies, Vol. 65, No. 3. (Jul., 1998), pp. 361-393

> Meyer, R., and J. Yu (2000), “Bugs for a Bayesian Analysis of Stochastic Volatility Models“, Econometrics Journal, vol. 3, pp 198-215

> Meyer, R., D. A. Fournier, and A. Berg (2003), “Stochastic volatility Bayesian computation using automatic differentiation and the extended Kalman filter“, Econometrics Journal (2003), volume 6, pp. 407–419.

> Neal, R. M. (2000), “Markov Chain Sampling Methods for Dirichlet Process Mixture Models”, Journal of Computational and Graphical Statistics,vol.9, No.2 (Jun., 2000), pp 249-265

> Petris, G., and S. Petrone (2011), “State Space Models in R”, Journal of Statistical Software, vol. 41, 4, May 2011

> Poon, S., and C. W. Granger (2003),“Forecasting volatility in financial markets: A review”, Journal of Economics Literature, 41, 478–539.

> Shepard, N., (2005), “Stochastic Volatility – Selected Readings”, Oxford University Press

> Taylor, S. J. (1994), “Modeling Stochastic Volatility: A Review and Comparative Study“, Mathematical Finance, Vol. 4, no. 2, (April 1994), pp. 183 - 204

Thank You