Mcmc Application To Stochastic Volatility Model

Using INR Exchange Data

Ankur Gupta

Dissertation
Submitted in Partial Fulfillment
Final year Project
5 Year BS-MS Dual Degree Program

Thesis Advisory Committee
Supervisor
Prof. U. V. Naik
Nimbalkar

Department of Mathematics
Indian Institute of Science Education and Research Pune

April 2011

Certificate

This is to certify that the work done by Ankur Gupta entitled "**MCMC Application To Stochastic Volatility Model**" has been carried out under my supervision and that it has not been submitted elsewhere for the award of any degree.

Signature of Student
Ankur Gupta

Signature of Supervisor
Prof. U.V.Naik Nimbalkar

Acknowledgements

I would like to thank my project guide **Prof.U.V.Naik Nimbalkar** for her continuous help in my project. She always listened to me and gave her advice to clarify my doubts.

Abstract

The main objective of my Masters thesis is to study the Dynamic structure that return prices or exchange rates exhibits. Intensive studies have been conducted for inferring the parameters and missing information. [1] paper studies this for ASEAN (Association of South East Asian Nations) markets. To study this I have chosen SV model as it mimics most of the stylized facts that exchange rates show. Also I have applied a Bayesian computation approach to infer the parameters associated with the SV model. I have chosen MCMC technique as the main approach since it solves some rigorous calculation issues which other techniques cannot overcome. Our data series includes exchange rates of Indian Rupees (INR) with United States Dollars (USD), and the period covers the crises of the ASEAN markets in 1997. Most of the part of thesis includes understanding of the basic and advanced concepts involved while applying MCMC technique to SV model. I have also studied MCMC application to some other models, which I have not mentioned in this thesis, like Geometric Brownian model etc, but I have concentrated my studies on SV model for the application of MCMC. Finally I have estimated the parameters involved in SV model by producing results from my own written MATLAB codes. The results produced are quite expected because of high level of persistence involved in the data. Also, the properties of MCMC can be quite easily visible from the graphs shown in the results section. It would be interesting for further research to come up with more time-effective codes.

Contents

Chapter 1

INTRODUCTION

Finding out the information about the state variables and the parameters from the observed asset prices is done through the empirical analysis of dynamic asset pricing models. Through this we can find the distribution of the parameters, θ, and state variables, X, conditional on observed prices, Y, which we denote by p(θ,X|Y). This distribution can be completely summarized by the marginal distributions p(θ|Y) and p(X|Y), i.e we can completely summarize parameter estimation and state variable estimation and also provides specification diagnostics.
Stochastic volatility model has been observed to be the most important model among other several financial time series models since it uses commonly observed change in the variance of the observed exchange rates as time evolves. The most important parameter in financial forecasting is Volatility as it measures the amount of fluctuation and randomness associated with price changes in asset pricing. Volatility gains over other parameters in a way as its magnitude can be captured via the movement in prices which can be best used for forecasting exchange rate or price in trading market. The other advantage is it is considered as a risk parameter in many asset return models, options and other derivative security pricing models.
Our aim in the thesis is to introduce a discrete SV model for obtaining a better estimate of the changing variance for the financial price returns. It is advantageous to use SV model over other models for forecasting the return prices. To achieve such a goal, modeling and inference must be carried out efficiently with the available price returns. For this Markov Chain Monte Carlo technique has been used.
Let us consider an example, in which a model of an equity price, S$_t$,whose variance, V$_t$, follows a square root process

$$\frac{dS_t}{S_t} = (r_t + \eta_v V_t)dt + \sqrt{V_t}dW_t^s$$

$$dV_t = \kappa(\theta_v - V_t)dt + \sigma_v\sqrt{V_t}dW_t^v$$

where W_t^s and W_t^v are two scalar Brownian motions and r_t is the instantanious spot interest rate. The goal in this model is to learn about volatility assets, using empirical asset pricing, $V = \{V_t\}_{t=1}^T$, the parameters that drive the evolution of the volatility processes, κ_v, θ_v, σ_v, and the risk premium, η_v, and assess model specification from observed prices, $Y = \{S_t\}_{t=1}^T$. The distributions $p(\kappa_v, \theta_v, \sigma_v, \eta_v|Y)$ and

$p(V|Y)$, which are marginals of $p(\kappa_v, \theta_v, \sigma_v, \eta_v, V|Y)$, determine parameter inference and estimate volatility respectively.

Generally, In continuous time models Characterizing $p(\theta, X|Y)$ is difficult for many reasons. First, the observed data are discretely picked while according to the models specification, asset prices and state variables should evolve continuously through time. Second, in most interesting models, there are state variables which are latent from the perspective of an econometrician. Third, practically the dimension of $p(\theta, X|Y)$ is very high dimension due to the state vector dimension. Fourth, mostly the distributions generated by most of the models for prices and state variables are non-normal and non-standard(e.g., models with stochastic volatility or jumps). Finally, in term structure and option pricing models, parameters forms are nonlinear or even in a non-analytic as the implicit solution to ordinary or partial differential equation.

The next chapter provides a description of MCMC methods and describes how they overcome these difficulties to generate samples from $p(\theta, X|Y)$. We provide a general description of the Bayesian approach to estimation, a detailed discussion of the components of MCMC algorithms and their convergence properties, then in further next chapter we provide the general details of stochastic volatility model , finally, we show how to estimate a number of parameter in stochastic volatility model with MCMC methods. Then in results section we will show the convergence of SV parameters μ, ϕ, σ^2 using the Indian FX data series.

1.1 Literature survey

The research, practice and implication of Markov Chain Monte Carlo(MCMC) has been increased a lot in recent years. From the huge volume of literature, some of them are [2] , [3] , [4] and [5] . All these papers and books provides a general review of the characteristics of the MCMC sampling.

Many researchers has suggested MCMC for carrying out Bayesian inference for SV models. for e.g. [6] and [7] were the first to apply MCMC efficiently for the SV model. Later much more contribution towards Bayesian computation of the MCMC for SV model has been given in research papers and further books like [8] , [9] and [10] . A good introduction and explanation of the MCMC approach is given in [3] , and is outlined in the next chapter.

Chapter 2

BAYESIAN COMPUTATION

Nowadays there are immense material to read about the basic and advanced concepts of Bayesian computation. I have chosen SV model for the application of Bayesian computation basically MCMC. [2] and [3] has explained these concepts in a very detailed manner. Most of the concepts of MCMC has been explained in an algorithmic way which I am going to explain in a detail manner in this chapter. [11] in his paper explained that it is not easy to obtain an explicit expression for the likelihood, so implicit specification of returns distribution is easy rather than explicit specification.

We can easily implement Bayesian approach. It involves a parameter θ which is treated like a random variable having its distribution over a parameter space. First we take the prior distribution of this θ without seeing any data. This prior distribution basically will tell us about our degree of belief for θ. Now in next step we will take point-wise data from the data series and apply Bayesian calculus to our prior distribution, this we'll continue till the last data point of data series and at each step update prior distribution which will finally become the posterior distribution. So, after taking into account the last data point, the prior becomes the posterior distribution. Below is the bayesian split of posterior distribution in terms of prior and likelihood :

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}$$

When we have posterior distribution in our hand we can do all inferences calculation like moment calculation, decision making, estimation etc. But to do this we have to evaluate the integral in the denominator which is tough to evaluate. To do this we apply a higher approach named Monte Carlo Markov Chain (MCMC), which is explained from next section onwards.

2.1 MCMC Method

In MCMC method, first we construct a Markov chain of sequence of random variables of k-dimension, $\{\theta_t, t = 1, 2, ..., k\}$. Depending upon the model, each variable θ_{t+1} is sampled from one step ahead conditional distribution $p(\theta_{t+1}|\theta_t)$. This one step ahead distribution is dependent only on the current state of the chain, θ_t. Due to regularity condition the Markov Chain has the property to attain an equilibrium or stationary distribution, which plays a central role in MCMC. By this method we can sample from any posterior distribution which can define the stationary distribution of the Markov Chain. So the key aspect in this method includes forming the one step ahead conditional distribution such that the stationary distribution of this chain will define the required Bayesian posterior distribution.

Now following the book [3] . Let D denote the observed data, and θ denote model parameters and missing data. Setting up a joint probability distribution $p(D, \theta)$ is required for formal inferencing over all random quantities. This joint distribution includes two parts: a prior distribution $p(\theta)$ and a likelihood $p(D|\theta)$. Specifying $p(\theta)$ and $p(D|\theta)$ provides a full probability model, in which

$p(D, \theta) = p(D|\theta)p(\theta)$

By observing D and applying Bayes theorem, we can determine the distribution of θ conditional on D.

$$p(\theta|D) = \frac{p(\theta)p(D|\theta)}{\int p(\theta)p(D|\theta)d\theta}$$

Above is the posterior distribution of θ, and is used for all Bayesian inference.

All the features of the posterior distribution are justifiable for Bayesian inference like calculating moments, quantiles, highest posterior density regions, etc. Since these quantities are expressed in terms of posterior expectations of functions of θ, the posterior expectation of a function f(θ) is

$$E[f(\theta)|D] = \frac{f(\theta)p(\theta)p(D|\theta)d\theta}{\int p(\theta)p(D|\theta)d\theta}$$

But in high dimension the integral in the denominator has been source of most difficult in Bayesian Inferencing. An alternative approach like numerical evaluation is used to evaluate, but for dimensions greater than 20 it is difficult to handle and mostly gives an inaccurate results. So, we use Monte Carlo approach, MCMC to evaluate this kind of integral which does not depends upon the dimension.

2.1.1 Calculating Expectations

To avoid involving an unnecessary Bayesian touch in the above problem let us restate above problem in more general terms. Let X be a vector of k random variables, with distribution $\pi(.)$. In Bayesian inference X will comprise of model parameters and missing data and $\pi(.)$ is the posterior distribution. Now modified problem involves evaluating the expectation

$$E[f(X)] = \frac{f(x)\pi(x)dx}{\int \pi(x)dx}$$

Let us allow that the distribution of X is known only upto a constant of normalization. That is, $\int \pi(x)dx$ is unknown which is a common situation in practice, for eg. in Bayesian inference, we know $p(\theta|D)\alpha p(\theta)p(D|\theta)$, but we cannot easily evaluate the normalization constant $\int p(\theta)p(D|\theta)d\theta$

2.1.2 Monte Carlo integration

The main problem that Monte Carlo integration solves is that it evaluates $E[f(X)]$ by drawing samples $X_t, t = 1, 2, ..., n$ from $\pi(.)$ and then approximating

$$[f(X)] \approx \frac{1}{n}\Sigma_{t=1}^{n} f(X_t)$$

when samples X_t are independent, law of large numbers ensure that the approximation can be made as accurate as desired by increasing the sample size n. Note that here n is under the control of the us: it is not the size of a fixed data sample.
In general , drawing samples X_t independently from $\pi(.)$ is not feasible, since $\pi(.)$ can be quite non standard. However, the X_t need not necessarily be independent. The X_t can be generated by any process which, draws samples throughout the support of $\pi(.)$ in correct proportions. One way of doing this is through a Markov Chain having $\pi(.)$ as its stationary dist. This is then Markov Chain Monte Carlo.

2.1.3 Markov Chains

Suppose we generate a sequence of random variables, $X_0, X_1, ...,$ such that at each time $t \geq 0$, the next state X_{t+1} is sampled from the distribution $p(X_{t+1}|X_t)$ which depends only on the current state of the chain, X_t. That is, given X_t, the next state X_{t+1} does not depend further on the history of the chain $X_0, X_1, ..., X_{t-1}$. This sequence is called a Markov chain.
How does the starting state X_0 affect X_t ? This question concerns the distribution of X_t given X_0, $p^{(t)}(X_t|X_0)$ will eventually converge to a unique stationary (or invariant or equilibrium) distribution, which does not depend on t or X_0. Let the stationary distribution is denoted by $\phi(.)$. Thus as t increases, the sampled points X_t will look increasingly like dependent samples from $\phi(.)$. Thus after a sufficiently long burn-in of say m iterations, points $X_t; t = m + 1, ..., n$ will be dependent samples approximately from $\phi(.)$. We can now use the output from the Markov Chain to estimate the expectation $E[f(X)]$, where X has distribution $\phi(.)$.
So,

$$\overline{f} = \frac{1}{n-m}\Sigma_{t=m+1}^{n} f(X_t)$$

2.2 The Metropolis-Hastings algorithm

The above equation shows how a Markov Chain can be used to estimate $E[f(X)]$, where the expectation is taken over its stationary distribution $\phi(.)$. This actually provides the solution to our problem but first we need to understand how to construct a Markov chain such that its stationary distribution $\phi(.)$ is precisely our distribution of interest $\pi(.)$. This can be done using Metropolis-Hastings algorithm. For the Metropolis -Hastings algorithm, at each time t, the next state X_{t+1} is chosen by first sampling a candidate point Y from a proposal distribution $q(.|X_t)$. Note that the proposal distribution may depend on the current point X_t. For eg., $q(.|X_t)$ might be a multivariate normal distribution with mean X and a fixed covariance matrix. The candidate point Y is then accepted with probability $\alpha(X_t, Y)$ where

$$\alpha(X, Y) = \min(1, \frac{\pi(Y)q(X|Y)}{\pi(X)q(Y|X)})$$

If the candidate point is accepted, the next state becomes $X_{t+1} = Y$. If the candidate is rejected, the chain does not move, i.e $X_{t+1} = X_t$. Thus Metropolis-Hastings algorithm is easy to operate.

Initiate X_0; set $t = 0$
Repeate{
Sample a point Y from $q(.|X_t)$
Sample a Uniform (0,1) random variable U
If

$$U \leq \alpha(X_t, Y) set X_{t+1} = Y$$

otherwise set

$$X_{t+1} = X_t$$

Increment t
}
Remarkably, the proposal distribution $q(.|.)$ can have any form and the stationary distribution of the chain will be $\pi(.)$. This can be seen from the following argument. The transition Kernel for the M-H algorithm is

$$p(X_{t+1}|X_t) = q(X_{t+1}|X_t)\alpha(X_t, X_{t+1}) + I(X_{t+1} = X_t)[1 - \int q(Y|X_t)\alpha(X_t, Y)dY],$$

The first term arises from acceptance of a candidate $Y = X_{t+1}$, and the second term arises from rejection, for all possible candidates Y. Using the fact that

$$\pi(X_t)q(X_{t+1}|X_t)\alpha(X_t, X_{t+1}) = \pi(X_{t+1}))q(X_t|X_{t+1})\alpha(X_{t+1}, X_t)$$

So, we obtain a detailed balance equation:

$$\pi(X_t)p(X_{t+1}|X_t) = \pi(X_{t+1})p(X_t|X_{t+1})$$

Integrating both side w.r.t X_t gives:

$$\int \pi(X_t p(X_{t+1}|X_t)dX_t = \pi(X_{t+1})$$

The L.H.S of the above equation gives the marginal distribution of X_{t+1} under the assumption that X_t is from the distribution $\pi(.)$. Therefore, if X_t is from $\pi(.)$, then X_{t+1} will also be. Thus once the sample from the stationary distribution has been obtained, all subsequent samples will be from that distribution. This only proves that the stationary distribution is $\pi(.)$. But for the complete justification of M-H algorithm we need to prove that $p^{(t)}(X_t|X_0)$ will converge to the stationary distribution. This can be proved by series of definitions and theorems which is there in appendix.

Chapter 3

STOCHASTIC VOLATILITY MODEL

An introduction of the discrete Stochastic Volatility model is needed in my work so as to achieve a better estimate of the changing variance that are found for the Asian financial price returns, so as to enable the automatic forecast at any given point of time. SV model of late have come into common use as they are easier to figure out and can also shed light on the matter of the stylized facts of volatility.
In financial time series, a certain stylized facts about volatility which is evident in price returns. These features noted in [12] and [13], play a crucial role in model construction and selection. I try to model a specification in which the following stylized features can be mimicked.
(1) Leptokurtic being the most common feature of volatility has long been recognized as the distribution of any financial returns. It has been reported by [14] that heavy tails of the distribution of observed returns indicates heavier tails than normal of the return distribution.
(2)The next stylized or realistic feature being recognized for financial price return is leverage or asymmetric effect which was coined by [15]. The term leverage corresponds to the negative correlation between stock price movements and volatility. We know that there is unexpected increase and decrease in the price, this thing was verified by [16] report for an asymmetric influence of positive and negative returns on the subsequent volatility. In his report he finds that bad news increases volatility more than good news.
(3)Next fact which the financial return prices reflects is persistence or volatility clustering, which implies that large movements in price. This is simple correlation of volatility.
(4)The final stylized fact is the correlation which is reflected by the interactive movement between markets, or sectors, or stocks in a sector, or indices, or currency exchange rates. We can easily see the volatilities of different financial markets move together, as a result of formal linkage or common factors amongst them.

3.1 Model Description

The SV model has been extensively used in quantitative finance since the early 1970's. This model does not only focus on volatility modeling with the help of observations that are made easier for the time varying variance in ARCH models, but also let the variance to have a latent stochastic structure. [17] in his early work on the microstructure of financial markets has taken stochastic volatility model to represent the random and uneven information vested in the financial market structure. Clake was the one, who not only did the work described in above lines but he also suggested that for the continuous time model of a log asset price, the SV model precisely represented an Euler motion. Also, option pricing tool was developed by [18] using this idea of Clake. [18] also suggested a positive diffusion process for asset prices with volatility. Since SV model has been able to explain the random and uneven flow of information , which is used in option pricing, this option pricing has become the most popular research areas in finance.

3.1.1 AR(1) Stochastic Volatility Model

Following [8], the simplest discrete time SV model can be defined by

$$Y_t = \epsilon_t \sigma_t, \tag{3.1}$$

where Y_t is the average corrected return on the asset price at time t, (t=1,...,n for all n$\geq$1). σ_t^2 follows the first order autoregressive (AR(1)) process. ϵ_t is assumed to be a series of independent, identically distributed random disturbances where $\epsilon_t \sim N(0,1)$. Consider the parameters h_t where

$$Y_t = \epsilon_t exp(h_t/2), \tag{3.2}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \tag{3.3}$$

ϵ_t and $\eta_t \sim N(0,\sigma^2)$ assumed to be independent normal white noise random processes. The parameters h_t are the log-volatility parameters, where
$h_t = log\sigma_t^2$
is the conditional variance at time t, and
$h_t|h_{t-1} \sim N(\mu + \phi(h_{t-1} - \mu), \sigma^2)$.
Parameters μ is a constant scaling factor, and parameter ϕ is the persistence parameter in this model. Parameter σ^2 is the conditional variance in the autoregressive log-volatility sequence. Equation 3.2 gives the so-called Log-Normal SV model. It can be shown that, by stationary (constancy and finiteness of the first two moments, and the autocovariance), the implied model for the initial state(and the implied marginal model for general t) is

$$h_1 \sim N(\mu, \frac{\sigma^2}{1-\phi^2}) \tag{3.4}$$

Squaring and logging equation 3.4 gives logY_t^2 which then has a linear, non-Gaussian state space representation

$$logY_t^2 = h_t + log\epsilon_t^2. \tag{3.5}$$

The one-step-ahead forecast density of the SV model can be defined as

$$Y_t|h_t \sim N(0, exp(h_t)), \tag{3.6}$$

where $t = 1, ..., n$. Following [8], the one step ahead predictive can also be approximated in the Monte Carlo sense as

$$f(y_t|y_{t-1}) \approx \frac{1}{N}\sum_{i=1}^{N} f(y_t|h_t^{(i)}, y_{t-1}), \tag{3.7}$$

where the state parameter $h_t^{(i)}$ $i = 1, ..., N$ is being sampled from the appropriate (posterior) probability distribution.

3.2 Properties of SV Model

[11] and [12] give a good review of the properties and also commonly and also commonly used estimation procedures for the SV models. The usual assumption of SV model is to assume $|\phi| < 1$ as this implies second-order stationarity. This gives the marginal form implied by 3.4 where

$$\mu_{h_t} = E[h_t] = \frac{\mu}{1-\phi}, \tag{3.8}$$

$$\sigma_{h_t}^2 = Var[h_t] = \frac{\sigma^2}{1-\phi^2} \tag{3.9}$$

where h_1 is drawn from the stationary distribution, h_t is a stationary process, and Y_t is stationary since ϵ_t is always stationary. From [11], considering the properties of the log-normal distribution, if h_t is stationary and r is even, there exist all the moments of Y_t. That is

$$E[Y_t^r] = E[\epsilon_t^r]E[exp(\frac{rh_t}{2})] = \frac{r!exp(\frac{r\mu_{h_t}}{2} + \frac{r^2\sigma_{h_t}^2}{8})}{2^{\frac{r}{2}}(\frac{r}{2})!} \tag{3.10}$$

when r is odd , all moments are zero, the kurtosis for the discrete SV model can be defined using the formula from 3.10 as

$$\kappa = \frac{E[Y_t^4]}{(E[Y_T^2])^2} = \frac{3exp(2\mu_{h_t} + 2\sigma_{h_t}^2)}{\{exp(\mu_{h_t} + (\frac{\sigma_{h_t}^2}{2}))\}^2} = 3exp(\sigma_{h_t}^2), \tag{3.11}$$

where $\kappa \geq 3$. The kurtosis calculation shows that the marginal SV model has heavier tails than an equivalent correlated Gaussian process. This is effectively because we are considering a scale-mixture of Gaussian densities.

This model captures the correlation between successive variances. The ACF between squared observations is defined as

$$\rho_{y_t^2}(s) = \frac{Cov[Y_t^2 Y_{t-s}^2]}{Var[Y_t^2]} = \frac{exp(\sigma_{h_t}^2\phi_1^s) - 1}{3exp(\sigma_{h_t^2}) - 1} \approx \frac{exp(\sigma_{h_t}^2) - 1}{3exp(\sigma_{h_t^2}) - 1}\phi^s = \frac{(\kappa/3) - 1}{\kappa - 1}\phi^s \tag{3.12}$$

for $s \geq 1$. In particular when ϕ is close to 1, there is a high degree of persistence in the volatility. The dynamic properties of the SV model can be presented using logs as shown in 3.5. [11] noted that $\log y_t^2$ follows an ARMA(1,1) model. If $\epsilon_t \sim N(0,1)$, then

$$E[\log \epsilon_t^2] = -1.2704, Var[\log \epsilon_t^2] = 4.93.$$

This is also used in [8], where the ACF of $\log y_t^2$ at lag s is approximately

$$\frac{\phi^s}{(1 + \frac{4.93}{\sigma_h^2})}$$

3.3 Likelihood function for SV Model

The full likelihood function under our discrete AR(1) SV model from section 3.1.1 is given by
$f(y, h|\mu, \phi, \sigma^2) = f(y|h, \mu, \phi, \sigma^2)f(h|\mu, \phi, \sigma^2)$.

$$f(y|h, \mu, \phi, \sigma^2) = \prod_{t=1}^{n} (\frac{1}{2\Pi})^{\frac{1}{2}} exp(\frac{-h_t}{2}) exp(-\frac{y_t^2}{2exp(h_t)}). \tag{3.13}$$

$$f(h|\mu, \phi, \sigma^2) = (\frac{(1-\phi^2)}{2\pi\sigma^2})^{\frac{1}{2}} exp\{-\frac{(1-\phi^2)}{2\sigma^2}(h_1-\mu)^2\} \times \prod_{t=2}^{n} (\frac{1}{2\pi\sigma^2})^{\frac{1}{2}} exp(-\frac{1}{2\sigma^2}(h_t-\mu-\phi(h_{t-1}-\mu))^2) \tag{3.14}$$

[8] notes that the ML for the SV model can only be obtained using the most intensive computational methods. [19] introduces an alternative approach of Quasi Maximum Likelihood(QML) using Kalman Filter.
An alternative approach generally used by econometricians is Generalized methods of moments(GMM). [7] used GMM estimation but finds that it is not an efficient method because it does not support the high volatility persistence.

3.4 Bayesian Inference for SV Model

For the AR(1) SV model introduced above, the full posterior distribution can be written as

$$p(\mu, \phi, \sigma^2, h|y) = \frac{p(y|\mu, \phi, \sigma^2, h)p(\mu, \phi, \sigma^2, h)}{\int\int\int\int p(y|\mu, \phi, \sigma^2, h)p(\mu, \phi, \sigma^2, h) d\mu d\phi d\sigma^2 dh}$$

where $h = (h_1, ..., h_n)$ and $y = (y_1, ..., y_n)$.
A joint prior distribution can be specified as

$$p(\mu, \phi, \sigma^2) = p(\mu)p(\phi)p(\sigma^2)$$

where(following from [**?**],
$\mu \sim N(0, \sigma_\mu^2)$
$p(\mu) = (\frac{1}{2\pi\sigma_\mu^2})^{\frac{1}{2}} exp\{-\frac{\mu^2}{2\sigma_\mu^2}\}$,

$\sigma^2 \sim IGamma(\omega_{\sigma^2}, \zeta_{\sigma^2})$,
$p(\sigma^2) = (\frac{\zeta_{\sigma 2}}{\Gamma(\omega_{\sigma 2})})(\frac{1}{\sigma^2})^{\omega_{\sigma^2}+1} exp(-\frac{\zeta_{\sigma 2}}{\sigma^2})$,
$\frac{\phi+1}{2} \sim Beta(\omega_\phi, \zeta_\phi)$,
$p(\phi) = \frac{2\Gamma(\omega_\phi+\zeta_\phi)}{\Gamma(\omega_\phi)\Gamma(\zeta_\phi)}(1+\phi)^{\omega_\phi-1}(1-\phi)^{\zeta_\phi-1}$.
Again, following [?], we choose $\omega_{\sigma^2} = 5, \zeta_{\sigma^2} = .001 \times 5 = .05$ for σ^2, and $\omega_\phi = 20, \zeta_\phi = 1.5$ for ϕ. From their experience dealing with high frequency data for this model, they notice that the estimated values of ϕ often lie between 0.995 to 0.999.

3.5 Markov Chain Monte Carlo For the AR(1) SV Model

There are four parameters in the SV model. As given by [8], the required full conditionals are defined in this section. The full conditional of μ and the states h_t are identical under each parameterization. Let us denote

$$m_t = h_t - \mu, d_t = h_t - \phi h_{t-1}, \overline{d} = \frac{1}{n-1}\sum_{t=2}^{n} d_t, \tag{3.15}$$

and

$$M_1 = \sum_{t=2}^{n} m_{t-1}^2, M_2 = \frac{\sum_{t=2}^{n} m_{t-1}m_t}{\sum_{t=2}^{n} m_{t-1}^2}. \tag{3.16}$$

The full conditional posterior distributions needed for implementation of the MCMC approach for this model can be written as follows: If
$h_{(t)} = (h_1, ..., h_{t-1}, h_{t+1}, ..., h_n)$,
then
(i) For μ :
$p(\mu|\phi, \sigma^2, h, y) \sim N(m_\mu, \nu_\mu)$,
where
$m_\mu = \frac{(1+\phi)h_1+(n-1)\overline{d}}{(1+\phi)+(1-\phi)(n-1)}$, $\nu_\mu = (\frac{(n-1)(1-\phi)^2+(1-\phi)^2}{\sigma^2} + \frac{1}{\sigma_\mu^2})^{-1}$
(ii) For h_1 :
$p(h_1|h_{(1)}, \mu, \phi, \sigma^2, y) \alpha exp\{-\frac{1}{2}[h_1 + y_1^2 exp(-h_1) + \frac{h_1-\mu-\phi(h_2-\mu)^2}{\sigma^2}]\}$
(iii) For h_t :
$p(h_t|h_{(t)}, \mu, \phi, \sigma^2, y) \alpha exp\{-\frac{1}{2}[h_t + y_t^2 exp(-h_t) + \frac{(1+\phi^2)(h_t-a_t)^2}{\sigma^2}]\}$
where
$a_t = \frac{(1-\phi^2)\mu+\phi(h_{t-1}+h_{t+1})}{1+\phi^2}$
(iv) For h_n :
$p(h_n|h_{(n)}, \mu, \phi, \sigma^2, y) \alpha exp\{-\frac{1}{2}[h_n + y_n^2 exp(-h_n) + \frac{h_n-\mu-\phi(h_{n-1}-\mu)^2}{\sigma^2}]\}$
and finally in (v) and (vi) where (ϕ, σ^2) are either sampled directly, or formed as functions of the sampled parameters.
(v) For σ^2 :
$p(\sigma^2|\mu, \phi, h, y) \alpha IGamma(\frac{n}{2} + h_\sigma, \frac{(1-\phi^2)}{2} m_1^2 + \frac{1}{2}\sum_{t=2}^{n}(m_t - \phi m_{t-1})^2 + \zeta_\sigma)$
(vi) For ϕ :

$p(\phi|\mu,\sigma^2,h,y)\alpha(1+\phi)^{\frac{\omega_\phi+1}{2}-1}(1-\phi)^{\frac{\zeta_\phi+1}{2}-1}exp\{-\frac{S(\phi)}{2\sigma^2}\}$,
where
$S(\phi)=m_1^2(1-\phi^2)+M_1(\phi-M_2)^2$
Note that the conditional density for ϕ and $\{h_t\}$ are not available in closed form, and so will be sampled using MH algorithm.
I have written the code in MATLAB which is there in appendix section.

Chapter 4

RESULTS:GRAPHS AND TABLES

In this chapter, I put all the results that I get. There are some tables which includes the results already obtained for ASEAN data series, I have updated them with my results of INR data series, so that comparison will be easy. Table 4.1 and Figure 4.1 has been used for prior analysis while all the rest of tables and figures have been used for posterior analysis. All the discussion and conclusion part I have done in next chapter.

4.1 Tables and Plots

FX Currency	Code	Time Period Dec/96 - Mar/03	n	Kurtosis $\overbrace{K_{R_t}}$	Zero Returns
Indian Rupees	INR	12/12/96 - 30/03/03	2300	36.61	672
Thai Baht	THB	12/12/96 - 30/03/03	2300	94.33	320
Singaporian Dollar	SGD	12/12/96 - 30/03/03	2300	17.82	248
Japanese Yen	JPY	12/12/96 - 30/03/03	2300	11.90	88
Hong Kong Dollar	HKD	12/12/96 - 30/03/03	2300	314.20	516

Table 4.1: **Details for the 5 FX Data series giving details about the period, Kurtosis, and number of zero returns for each currency. The details of other four data series is taken from [1]**

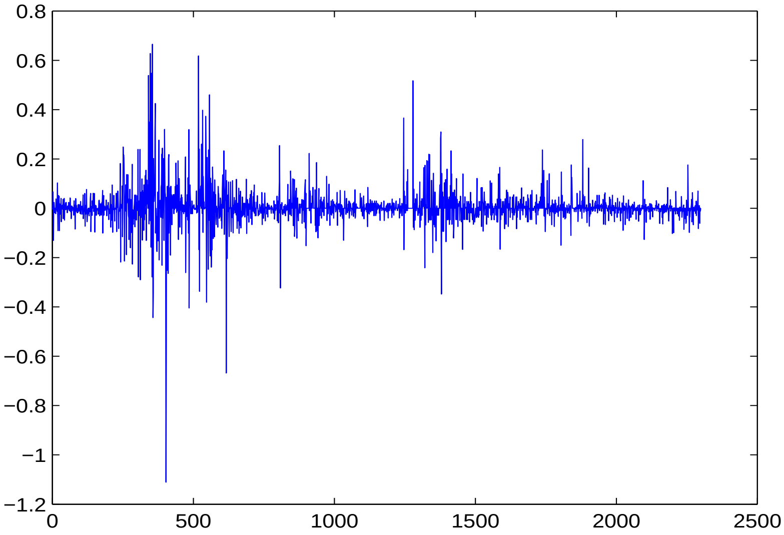

Figure 4.1: **Plot of daily INR return. X-axis:Day number, Y-axis: Value of return price**

SV Model Parameters	μ Mean,Median,Std	ϕ Mean,Median,Std	σ^2 Mean,Median,Std
INR	-1.2237, -1.2028, 0.4398	0.9273, 0.7908, 1.0638	2.0337, 2.0492, 1.0237
THB	-2.4985, -2.5023, 0.1825	0.8696, 0.8711, 0.0302	1.0735, 1.0374, 0.3079
SGD	-3.5635, -3.5633, 0.0909	0.5414, 0.5417, 0.0402	2.6956, 2.6813, 0.3168
JPY	-1.7500, -1.7502, 0.0602	0.3312, 0.3316, 0.0538	1.5748, 1.5719, 0.1695
HKD	-9.8943, -9.8965, 0.1481	0.7947, 0.7951, 0.0237	1.8675, 1.8644,0.1920

Table 4.2: **Posterior Statistics for SV model for INR data series. Statistics for ASEAN nations has been taken from [1]**

SV Model Parameters	μ Mean,Median,Std	ϕ Mean,Median,Std	σ^2 Mean,Median,Std
INR Ist series	-1.2237, -1.2028, 0.4398	0.9273, 0.7908, 1.0638	2.0337, 2.0492, 1.0237
INR Ist series	-1.2101, -1.2199, 0.4428	0.8980, 0.7780, 1.0440	1.9820, 1.9960, 1.0190

Table 4.3: **Posterior Statistics for two parallel INR Data series**

SV Model	κ Mean, Std	$rho_{y_t^2}$ Mean, Std	$rho_{\log y_t^2}$ Mean, Std
INR	42.6139, 540.9215	0.2757, 0.3196	0.2891, 0.3616
THB	276.9011, 169.9642	0.2871, 0.0096	0.4068, 0.0248
SGD	143.3969, 48.0058	0.1776, 0.0129	0.2358, 0.0176
JPY	17.9572, 3.2092	0.0969, 0.0158	0.0874, 0.0152
HKD	522.4745, 211.7979	0.2637, 0.0080	0.4033, 0.0218

Table 4.4: **Posterior Statistics for kurtosis, ACF and** $\log$**ACF**

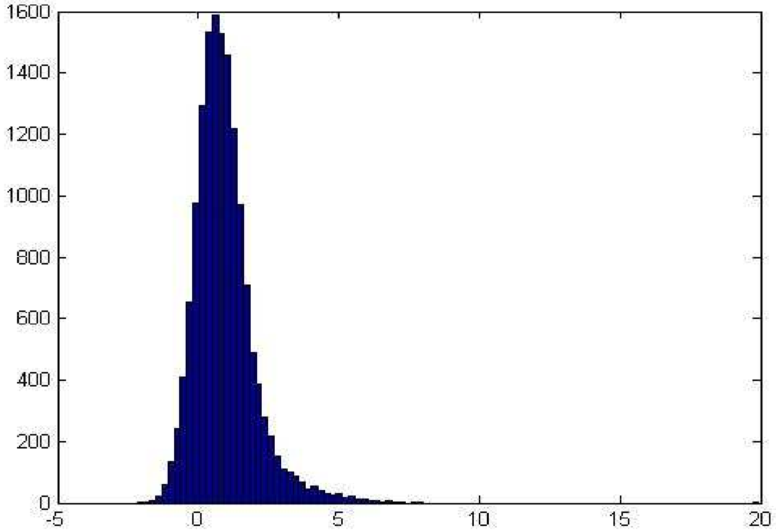

Figure 4.2: **Plot of SV Model for ϕ using INR Data series. X-axis: Value of ϕ, Y-axis: Number of times the corresponding X value attains**

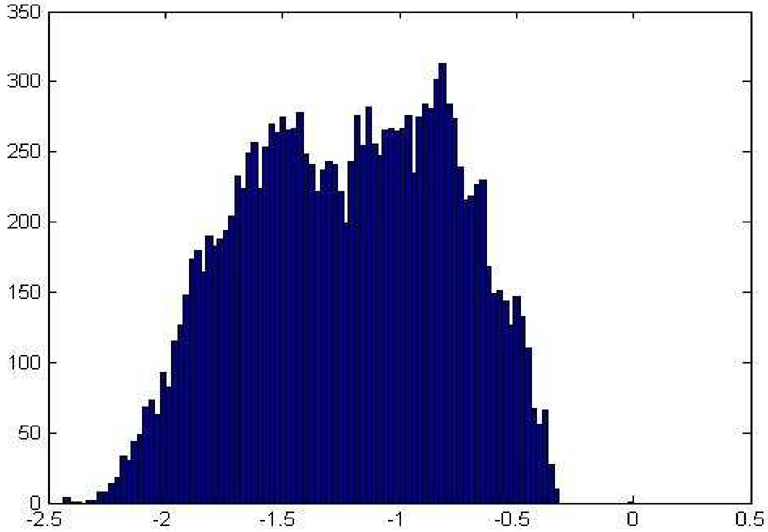

Figure 4.3: **Plot of SV Model for μ using INR Data series. X-axis: Value of μ, Y-axis: Number of times the corresponding X value attains**

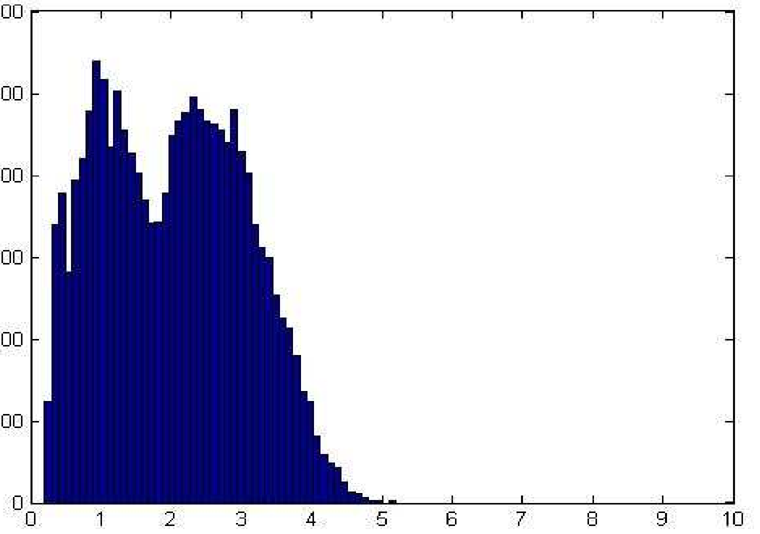

Figure 4.4: **Plot of SV Model for σ^2 using INR Data series. X-axis: Value of σ^2, Y-axis: Number of times the corresponding X value attains**

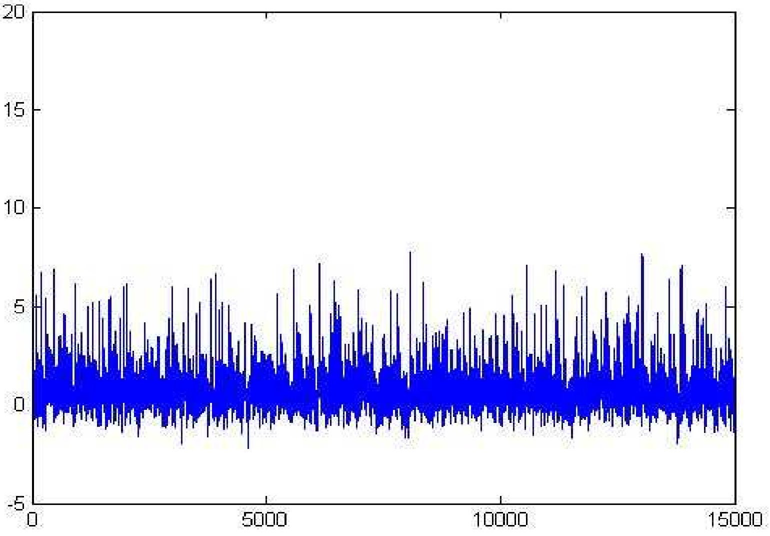

Figure 4.5: **Plot of ϕ time series using INR Data series. X-axis: iteration number, Y-axis: Value of ϕ**

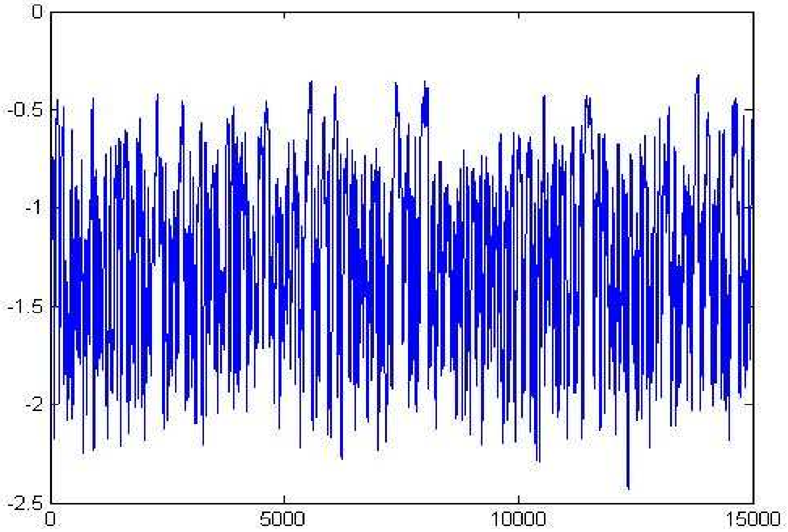

Figure 4.6: **Plot of μ time series using INR Data series. X-axis: iteration number, Y-axis: Value of μ**

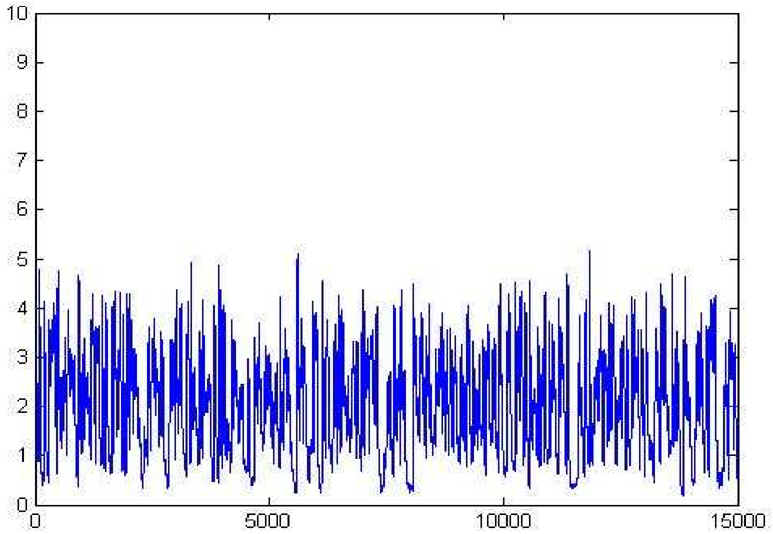

Figure 4.7: **Plot of σ^2 time series using INR Data series. X-axis: iteration number, Y-axis: Value of σ^2**

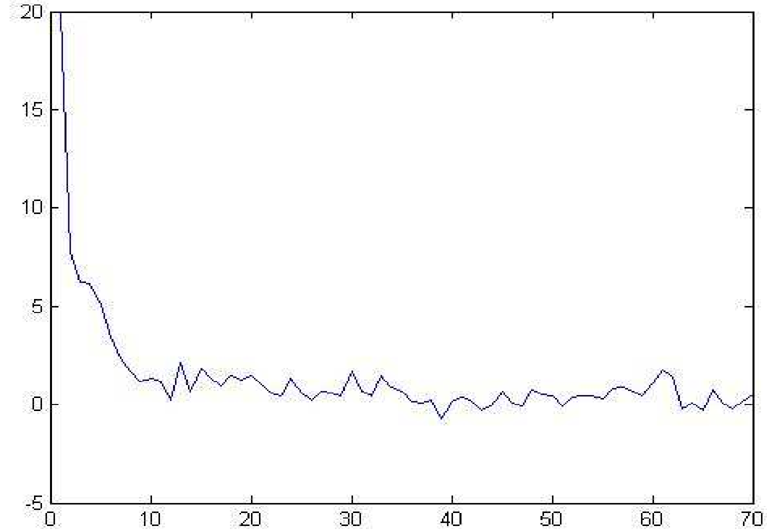

Figure 4.8: **Expansion of ϕ time series. X-axis: iteration number, Y-axis: Value of ϕ**

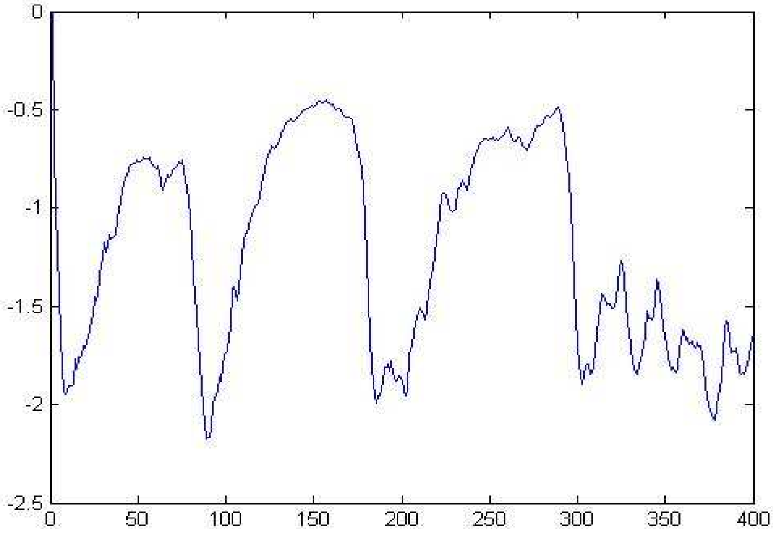

Figure 4.9: **Expansion of μ time series. X-axis: iteration number, Y-axis: Value of μ**

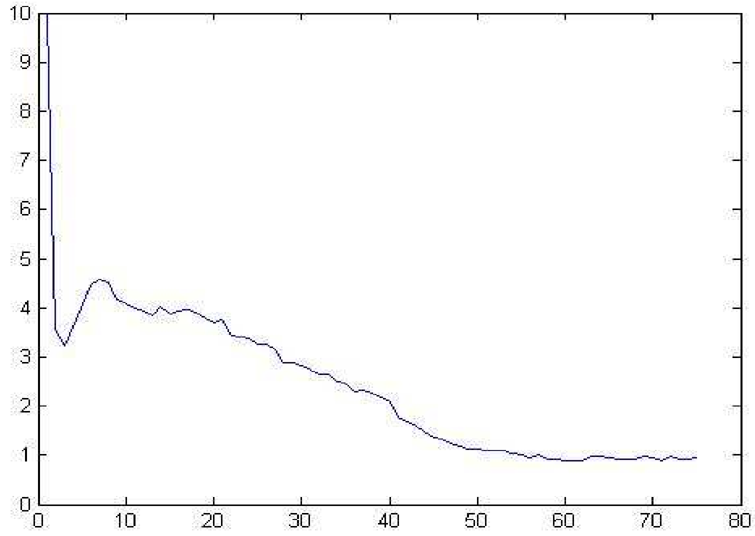

Figure 4.10: **Expansion of σ^2 time series. X-axis: iteration number, Y-axis: Value of σ^2**

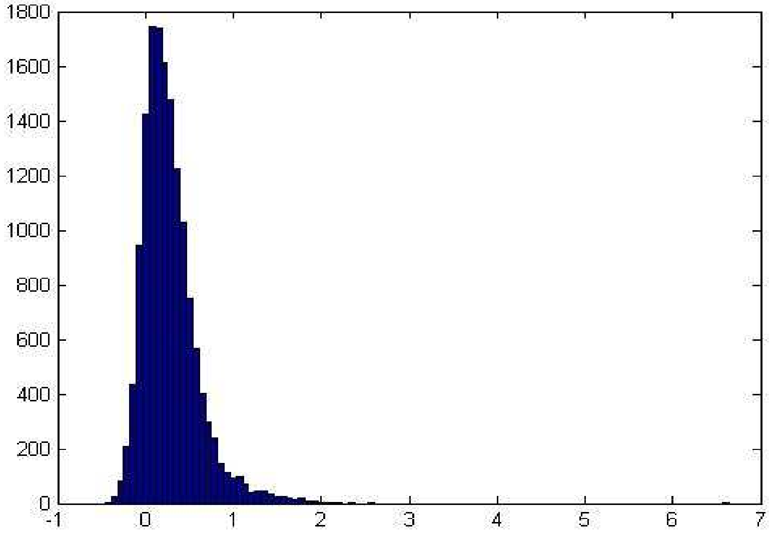

Figure 4.11: **ACF plot of y_t^2 for the INR Data series to assess autocorrelation. X-axis: ACF of y_t^2, Y-axis: Number of times the corresponding X value attains**

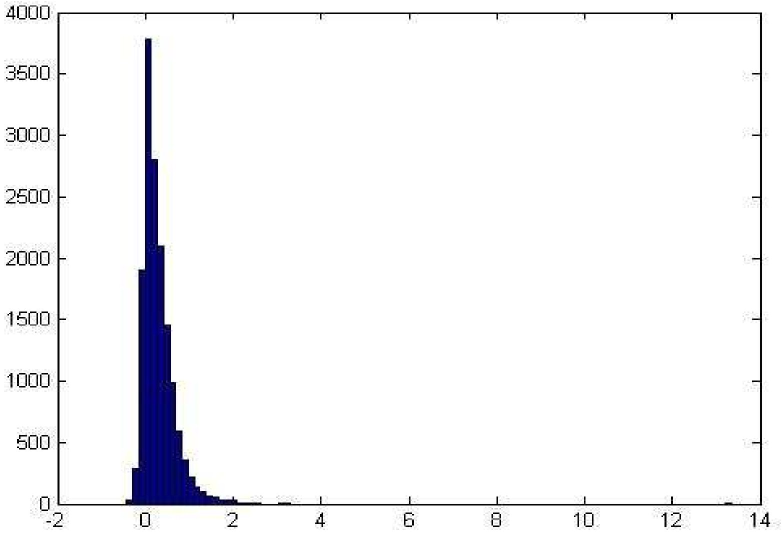

Figure 4.12: **ACF plot of $\log y_t^2$ for the INR Data series to assess autocorrelation. X-axis: ACF of $logy_t^2$, Y-axis: Number of times the corresponding X value attains**

Chapter 5

Discussion

In this chapter, I have done the exploratory and posterior data analysis of the results in *chapter 4* leading to some discussions. In the end, I have discussed some suggestions for further research in this topic.

5.1 Exploratory data analysis

The data series of Indian currency (INR) against US Dollar (USD) is obtained from Oslen and Associates (www.oanda.com). n=2300 data points are collected covering the period from the 12th Dec 1996 to 30th March 2003. Ihave chosen this period because some inferences are already in place for the same period for the ASEAN markets. So, it will be easier to compare my results with them.
Let $\{y_t\}$ denote the data points with $t = 1, 2, ..., n$. The Return R_t is defined as

$$R_t = 100 \times \log(\frac{y_t}{y_{t-1}})$$

which is also known as continuous compound or log returns. We know $\log(1+x) \rightarrow x$ for small values of x, so $\log(\frac{y_t}{y_{t-1}})$ approximates to $\frac{(y_t - y_{t-1})}{y_{t-1}}$ for small values of this relative change. The aim is to calculate the volatility or variability involved in $\{y_t\}$ and how this volatility evolves in time. I have calculated the kurtosis and zero return of the above approximation which is shown in *Table 1*. Also *Figure 1* represents the a priori data series.

5.2 Posterior data analysis

I have obtained the estimates of three parameters involved in the SV model $i.e$ $\{\mu, \phi, \sigma^2\}$ through MCMC sampling. The obtained results are shown in *figures 2,3,4*. But we can obtain the estimate only when the parameters converge, which I have checked by running multiple parallel chains with different starting values. *Table 3* shows that estimated statistics converge with different starting values. I have used Metropolis Hastings algorithm to sample states in MCMC sampler. The chain has been run for 20000 iterations with a burn-in of 5000 points, convergence

of statistics suggests that this much of iterations are enough.
The histogram and trace plots for the three parameters (μ, ϕ, σ^2) are shown in *Figure 2,3,4 and Figure 5,6,7*, and the estimated statistics are summarized in *Table 3* for posterior analysis. The trace plots *(Figure 5,6,7)* clearly demonstrates the convergence of the parameters.
To see any autocorrelation in the data, I have plotted ACF Plots in *Figure 10,11.* To compare the estimates of the posterior statistics like Mean, Median and Standard deviation of the parameters refer to *Table 2*. Finally these parameters are transformed to calculate the kurtosis and ACF of the series y_t^2 and $\log y_t^2$, see *Table 4*.

5.3 Discussion

The estimate of kurtosis can suggest one of the stylized facts that I have mentioned. I have obtained 42.61 as the mean value of Kurtosis which is a relatively moderate value as compared with the mean values of other data series. *Table 4* shows the moderate value of mean and median for kurtosis which implies that Indian currency returns distribution will have a moderate tail. So, the most persistent series appears to be moderately tailed with estimated mean value of kurtosis to be 42.61. If we compare the exploratory data value of kurtosis, i.e 36.61, with the estimated value of 42.61, there is not much difference, the value slightly increases, thereby signalling to use a slightly heavier tailed distribution like students t distribution.
Table 2 shows the posterior mean and median of ϕ. The obtained mean and median value, i.e 0.9273 and 0.7908, are comparitively larger than the mean and median values of the available data series from ASEAN nations. The high value of persistence implies that Indian currency has the most stable structure among all nations from *Table 2*. We were expecting this result a-priori since the number of zero returns, 672, was the highest from the data itself, which in turn also reflects the stability. On the other hand, the value of mean and median of σ^2 is on the higher side, which reflects that the Indian currency exchange cannot be easily predictable in terms of volatility.
Overall, we can say that the chosen SV model for applying MCMC sampler gives a good estimation for persistence factor but it is not a decent model for predicting the value of μ, since high zero returns should correspond to high negative mean value for μ, but we get low negative mean value.

5.4 Further suggestions

I also implement MCMC sampler on Geometric Brownian model for INR data series and found that SV model allows for a better estimate of the parameters in terms of comparing the predicted observed returns, but one needs to study some further model selection criteria.
Further, It remains to be seen whether MCMC sampler can be extended to other existing financial time series models to obtain a better estimate of the statistics.

Also, one can enhance the time-efficiency of the analysis by employing less time-consuming algorithms and codes, if possible.

References

[1] W. Surapaitoolkorn, Bayesian markov chain monte carlo(mcmc) for stochastic volatility model using fx data.

[2] B. J. M., A. Smith, Bayesian theory, John Wiley and Sons.

[3] R. S. Gilks, W.R., D. Spiegelhalter, Markov chain monte carlo in practice, Chapman and Hall.

[4] D. Garmerman, Markov chain monte carlo, Chapman and Hall.

[5] C. Roberts, G. Casella, Monte carlo statistical methods, New York: Springer-Verlag.

[6] N. Shephard, Fitting non-linear time series models, with applications to stochastic variance models, Journal of applied Econometrics 8 (1993) 135–152.

[7] P. N. Jacquier, E., P. Rossi, Bayesian analysis of stochastic models(with discussion), Journal of Business and Economic Statistics 12 (1994) 371–417.

[8] S. N. Kim, S., S. Chib, Stochastic volatility: Likelihood inference and comparison with arch models, Review of Economic Studies 65 (1998) 361–393.

[9] N. Shephard, M. Pitt, Likelihood analysis of non-gaussian measurement time series, Biometrika 84 (1997) 653–667.

[10] M. Pitt, N. Shephard, Analytic convergence rates and parameterization issues for the gibbs sampler applied to state space models, Journal of Time series Analysis 20 (1999) 63–85.

[11] N. Shephard, Statistical aspects of arch and stochastic volatility (1996) 1–55.

[12] H. A. Ghysels, E., E. Renault, Stochastic volatility, Hand book of statistics 14 (III) (1996) 209–235.

[13] A. B. Aydemir, Volatility modelling in finance, Forecasting Volatility in Financial Markets Chapter 1 (1998) 1–46.

[14] B. Mandelbrot, The variation of certain speculative prices, Journal of Business 36 (1963) 394–416.

[15] F. Black, Studies in stock price volatility changes, American Statistical Association (1976) 171–181.

[16] N. V. Engle, R.F., Measuring and testing the impact of news on volatility, Journal of Finance 48 (1993) 1749–1778.

[17] P.K.Clark, A subordinated stochastic process model with finite variances for speculative prices, Econometrica 41 (1973) 135–136.

[18] J. Hull, A. White, The pricing of options on assets with stochastic volatilities, Journal of Finance 42 (1987) 200–281.

[19] M. C. Romano, M. Thiel, J. Kurths, I. Z. Kiss, J. L. Hudson, Estimation of an asymmetric stochastic volatility model for asset returns, Journal of Business and Economic Statistics 14 (1996) 429–434.

Appendix A

Definitions and Theorems

Here I have written down some series of Definitions and Theorems which can prove that $p^{(t)}(X_t|X_0)$ will converge to the stationary distribution. These Definitions and Theorems have been taken from chapter 4 of the book [3]

A.1 Def 2.1

A Markov Chain is ϕ irreducible for a probability distribution ϕ on E if $\phi(A) > 0$ for a set $A \subset E$ implies that

$$P_x\{\tau_A < \infty\} > 0$$

for all $x \epsilon E$. A chain is irreducible if it is ϕ-irreducible, then ϕ is called an irreducibility distribution for the chain.

A.2 Def 2.2

An irreducible Markov chain with maximal irreducibility distribution ψ is recurrent if for any set $A \subset E$ with $\psi(A) > 0$ the conditions

(i) $P_x\{X_n \epsilon A$ infinitely often$\} > 0$ for all x

(ii)$P_x\{X_n \epsilon A$ infinitely often$\} = 1$ for ψ-almost all x

are both satisfied. An irreducible recurrent chain is positive recurrent if it has an invariant probability distribution otherwise it is null recurrent.

A.3 Thm 2.1

Suppose a Markov Chain $\{X_n\}$ is irreducible and has invariant distribution π. Then the chain is π irreducible, π is a maximal irreducibility distribution, π is the unique invariant distribution of the chain, and the chain is positive recurrent.

A.4 Thm 2.2

Suppose $\{X_n\}$ is an irreducible Markov Chain with transition kernel P and invariant distribution π. Define the average transition kernel $\overline{P^n}$ by

$$\overline{P^n}(x, A) = \frac{1}{n+1} \sum_{i=0}^{n} P^i(x, A)$$

for all $x \epsilon E$ and $A \subset E$. Then

$$||\overline{P^n}(x, .) - \pi(.)|| \to 0$$

for π-almost all x.

A.5 Thm 2.3

Suppose $\{X_n\}$ is an irreducible Markov Chain with transition kernel P and invariant distribution π, and let f be a real valued function on E such that $\pi|f| < \infty$. Then

$$p_x\{\overline{f^n} \to \pi f\} = 1$$

for π-almost all x, where $\overline{f^n}$ is given by

$$\overline{f^n} = \frac{1}{1+n} \sum_{i=1}^{n} f(X_i)$$

A.6 Thm 2.4

Suppose $\{X_n\}$ is an irreducible, aperiodic Markov chain with transition kernel P and invariant distribution π. Then

$$||P^n(x, .) - \pi(.)|| \to 0$$

for π-almost all x.

Appendix B

MATLAB Code for Bayesian Computation

I have written down a MATLAB code and used the Metropolis-Hastings algorithm for the Bayesian computation explained in Chapter-3.

B.1 Main code(mcforex.m)

```
clc
clear all
n=2299;
loop=15000;
phi=zeros(1,loop);
ss=zeros(1,loop);
h=ones(1,n,loop);
phi(1,1)=.997;
ss(1,1)=1;
mu=zeros(1,loop);
y=data;
sms=1;
m1=0;
m2=0;
vb=0;

   for i=1:loop-1
a=h(1,:,i);
mu(i+1)=mu1(phi(i),a,n,ss(i),sms);
[h(1,1,i+1),nu]=hone(y,mu(i+1),phi(i),a,ss(i),h(1,1,i));
a(1,1)=h(1,1,i+1);
for j=2:n-1
h(1,j,i+1)=ht(y(j),mu(i+1),phi(i),h(1,j-1,i+1),h(1,j+1,i+1),ss(i),h(1,j,i));
a(1,j)=h(1,j,i+1);
```

```
    end
h(1,n,i+1)=hn(n,y,mu(i+1),phi(i),a,ss(i),h(1,n,i));
a(1,n)=h(1,n,i+1);
[ss(i + 1), vb] = ss1(n, a, mu(i + 1), phi(i));
[phi(i + 1)] = phi1(n, a, mu(i + 1), ss(i + 1));

    end
```

B.2 function(mu.m)

```
function [pm]=mu1(phi,h,n,ss,sms)
d=h(1);
for i=2:n
dt=h(i)-phi*h(i-1);
d=d+dt;
end
db=d/(n-1);
mm=((1+phi)*h(1)+(n-1)db)/((1+phi)+(1-phi)*(n-1));
vm=(((n * (1 - phi)^2)/ss) + (1/sms))^-1;
pm=normrnd(mm,vm);
```

B.3 function(hone.m)

```
function [h1,h]=hone(y,mu,phi,h,ss,st)
alpha = 2.43;
k=0;
pdf = @(x)exp((-1/2) * (x + y(1)^2 * exp(-x) + ((x - mu - phi * (h(2) - mu))^2)/ss));
proppdf = @(x, y)gampdf(x, floor(alpha), floor(alpha)/alpha); proprnd = @(x)sum(exprnd(floor(
= 1;
h = mhsample(st,nsamples,'pdf',pdf,'proprnd', proprnd,'proppdf',proppdf);
h1=h;
```

B.4 function(ht.m)

```
function [ht]=ht(yt,mu,phi,htmo,htpo,ss,st)
at = ((1 - phi^2) * mu + phi * (htmo + htpo))/(1 + phi^2);
alpha = 2.43;
k=0;
pdf = @(x)exp((-1/2) * (x + yt^2 * exp(-x) + ((1 + phi^2) * (x - at)^2)/ss)); proppdf =
@(x, y)gampdf(x, floor(alpha), floor(alpha)/alpha); proprnd = @(x)sum(exprnd(floor(alpha)/alp
= 1;
ht1 = mhsample(st,nsamples,'pdf',pdf,'proprnd', proprnd,'proppdf',proppdf);
ht=ht1;
```

B.5 function(hn.m)

function [hn]=hn(n,y,mu,phi,h,ss,st)
alpha = 2.43;
k=0;
st=gamrnd(floor(alpha),floor(alpha)/alpha);
pdf = $@(x)exp((-1/2) * (x + y(n)^2 * exp(-x) + ((x - mu - phi * (h(n-1) - mu))^2)/ss))$; proppdf = $@(x,y)gampdf(x, floor(alpha), floor(alpha)/alpha)$; proprnd = $@(x)sum(exprnd(floor(alpha)/alpha, floor(alpha), 1))$;nsamples = 1;
hn1 = mhsample(st,nsamples,'pdf',pdf,'proprnd', proprnd,'proppdf',proppdf);
hn=hn1;

B.6 function(phi.m)

function [phi]=phi1(n,h,mu,ss)
alpha = 2.43;
m=zeros(1,n);
omphi=20;
ziphi=1.5;
for i=1:n
m(i)=h(i)-mu;
end
c=0;
d=0;
for i=2:n
$k = m(i-1)^2$;
c=c+k;
end
m1=c;
for i=2:n
l=m(i-1)*m(i);
d=d+l;
end
m2=d/m1;
pdf = $@(x)((1+x)^{(}((omphi+1)/2)-1))*((1-x)^{(}((ziphi+1)/2)-1))*exp(-(((m(1)^2)*(1-x^2)) + (m1*(x-m2)^2))/(2*ss))$;
proppdf = $@(x,y)gampdf(x, floor(alpha), floor(alpha)/alpha)$;
proprnd = $@(x)sum(exprnd(floor(alpha)/alpha, floor(alpha), 1))$;
nsamples = 1;
phi = mhsample(1,nsamples,'pdf',pdf,'proprnd', proprnd,'proppdf',proppdf);

B.7 function(ss.m)

```
function [ss]=ss1(n,h,mu,phi)
wss=5;
ziss=.005;
a=(n/2)+wss;
m=zeros(1,n);
for i=1:n
m(i)=h(i)-mu;
end
c=0;
for i=2:n
$k = (m(i) - phi * m(i-1))^2$;
c=c+k;
end
c=c/2;
b=$(((1 - phi^2) * m(1)^2)/2) + c + ziss$;
ss=1./gamrnd(a,b);
```