Journal of Data Science,17(3). P. 593 - 612,2019
DOI:10.6339/JDS.201907_17(3).0008

A Comparison between Bayesian and Frequentist methods in Financial Volatility with Applications to Foreign Exchange Rates

Steve S. Chung[1], Jalen Harris[1], Christopher Newmark[1], and Diana Yeung[2]

[1]*Department of Mathematics, California State University, Fresno*
[2]*Department of Mathematics, University of Notre Dame*

ABSTRACT

In this paper, a comparison is provided for volatility estimation in Bayesian and frequentist settings. We compare the predictive performance of these two approaches under the generalized autoregressive conditional heteroscedasticity (GARCH) model. Our results indicate that the frequentist estimation provides better predictive potential than the Bayesian approach. The finding is contrary to some of the work in this line of research. To illustrate our finding, we used the six major foreign exchange rate datasets.

Key words: Bayesian; Financial time series; Foreign exchange rates; Frequentist; Volatility

1. Introduction

In the last few decades, volatility in financial time series has been of a key interest to both academics and practitioners as uncertainty is at the heart of financial decisions. Volatility plays a critical role in pricing derivatives, calculating measures of risk, and hedging. Since the gold standard abandonment in 1971, asset prices and stock markets began to broadly change and searching for predictive volatility modeling has been one of the major areas in time series analysis. Early work on volatility includes the ARCH (autoregressive conditional heteroscedasticity) of Engle (1982) and the GARCH (generalized autoregressive conditional heteroscedasticity) of Bollerslev (1986), which have become the benchmark models for estimating the volatility. ARCH/GARCH and their extended implementations have been proven to be a successful tool in modeling the conditional variance of financial time series data. A few examples are as follow. Wang et al. (2010) investigated volatility on Shanghai Stock Exchange with high-frequency intraday data. Huang et al. (2012) investigated the performance of GARCH models in option pricing. More recently, Jahufer (2015) has used GARCH models to examine Sri Lanka stock market using non-parametric specification test.

The traditional frequentist approach uses the (conditional) maximum likelihood estimation (MLE) technique to estimate the parameters in the GARCH or GARCH-type models. We briefly describe this method in the next section and one can refer to Fan and Yao (2005) for more details. An- other technique that has gained momentum in recent years is the Bayesian approach, which takes into account prior information to estimate the posterior distribution. Nakatsuma (1999) developed three Bayesian methods: Markov chain Monte Carlo, Laplace approximation and quadrature formula to estimate the parameters of the ARMA-GARCH model. Bauwens (1998) explained how a Gibbs sampler can be implemented to perform the inferences on Bayesian GARCH models. Vrontos (2012) proposed a full Bayesian analysis of GARCH and Exponential-GARCH (EGARCH) model on parameter estimation, model selection, and volatility prediction.

The Bayesian method has been an alternative way to model datasets in many different fields. The comparison of GARCH models under frequentist and Bayesian has garnered some attention in research. Nakatsuma (1996) conducted a study which focuses on this comparison. Based on a small sample Monte Carlo experiment, they found that the Bayesian approach performs better than the frequentist approach when comparing the mean square errors of the posterior mean in the ARMA-GARCH models. Hoogerheide (2012) examined density prediction of stock index returns us- ing GARCH models under both frequentist and Bayesian estimation. They showed that there is no significant difference between the qualities of whole density forecast, while Bayesian estimation exhibits better left-tail fore- cast accuracy. More recently, Sigauke (2016) modeled the Johannesburg Stock Exchange (JSE) using the Bayesian and frequentist approaches and concluded the Bayesian Autoregressive Moving Average-Generalized Autoregressive Conditional Heteroskedasticity (BARMA-GARCH-t) provided a better fit for the data than the standard ARMA-GARCH-t model. In a more general setting, studies have been conducted to compare the Bayesian and frequentist methods. Wagenmakers et al. (2008) advocate the use of Bayesian inference in the field of psychology. Samaniego (2010) gives the comparison of the Bayesian and frequentist approaches to estimation. Albers et al. (2018) outline the ramifications of using frequentist and Bayesian analyses. In our work, we show that the traditional frequentist approach renders better predictive performance than the Bayesian approach.

The rest of the paper is organized as follows. Section 2 introduces the GARCH model along with the maximum likelihood estimation and Bayesian methodologies. Section 3 describes the results and Section 4 provides the discussion.

2. Methods

Let $\{x_t\colon t \in Z\}$ be a stochastic process that is adapted to filtration $\{F_t\colon t \in Z\}$, where $F_t = \sigma(\{x_s : s \le t\})$ and $\sigma(\{x_s\})$ is a sigma-field generated by $\{x_s\}$. Following Geweke (1993), we assume

$$x_t = \mu + \epsilon_t \left(\frac{v-2}{v} w_t \sigma_t^2\right)^{1/2} \quad t = 1, \ldots, T, \tag{1}$$

where ϵ_t are innovations and $\epsilon_t|F_{t-1}$ either follow a standard normal distribution or a t-distribution with v degrees of freedom. Although the mean, μ, can be time dependent in practice and modeled separately, we fix this value to be zero. In this work, we are primarily concerned with σ_t, the volatility, in time series economics. A plethora of works have been devoted to modeling this latent variable in the last thirty years and the work is still ongoing. As mentioned previously, the pioneer work on volatility is the ARCH/GARCH model of Engle (1982) and Bollerslev (1986). The GARCH model with order (1,1) (or GARCH(1,1)) assumes

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{t-1}^2 + \beta \sigma_{t-1}^2 \tag{2}$$

Our main focus is based on this GARCH(1,1) by examining the predictability of σ^2 under two cases for w_t: (1) fixed $w_t = v/(v-2)$ and (2) $w_t \sim$Inv-Gam($v/2$, $v/2$). The details are given in the following subsections.

In practice, if y_t is a stock price then the log-return series x_t is defined as

$$x_t = \log(y_t) - \log(y_{t-1})$$

This measures the relative changes in the stock price. The above form can also be written as:

$$\log(y_t) - \log(y_{t-1}) = \log\left(1 + \frac{y_t - y_{t-1}}{y_{t-1}}\right) \approx \frac{y_t - y_{t-1}}{y_{t-1}}$$

Many financial studies use the return series xt instead of price series yt for many benefits. First, the returns are scale-free. Second, they have more attractive statistical properties than the price series and third, they are time-additive. The reader can refer to Tsay (2010) for more details and elaboration.

2.1 Frequentist GARCH Estimation

In the traditional frequentist statistics, the parameters are fixed unknown constants. Under this framework, we fix $w_t = v/(v\text{-}2)$ so that the equation (1) becomes

$$x_t = \mu + \epsilon_t \sigma_t, \tag{3}$$

where ϵ_t follow N (0, 1) or t_v. Under a standard normal distribution, the likelihood function of $x = (x_1, ..., x_T)^T$ is defined as

$$L(\alpha_0, \alpha_1, \beta \mid x) = \frac{1}{(2\pi)^{\frac{T}{2}}} \prod_{t=p+1}^{T} [\sigma_t^2]^{-\frac{1}{2}} \exp\left[-\sum_{t=p+1}^{T} \frac{x_t^2}{2\sigma_t^2}\right] \tag{4}$$

and under a t-distribution with v degrees of freedom, the likelihood function is defined as

$$L(\alpha_0, \alpha_1, \beta \mid x) = \prod_{t=p+1}^{T} \left[\frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi\sigma_t^2}\,\Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{x_t^2}{v\sigma_t^2}\right)^{-(v+1)/2}\right] \tag{5}$$

The maximum likelihood (ML) estimators are the maximizers of the functions above. Note that σ_t^2 is a function of the unknown parameters α_0, α_1,and β and it depends on the past squared return series and the past squared volatility σ_t^2. In addition, the likelihood is conditioned on $(x_1^2, x_2^2, \ldots, x_p^2)$ and $(\sigma_1^2, \sigma_2^2, \ldots, \sigma_p^2)$. The reader is referred to Fan and Yao (2005) for more details. In our work, we used the nonlinear optimization under the augmented Lagrange method which is implemented in the R package ***solvnp*** of Ghalanos (2011) in rugarch of Ghalanos (2016).

2.2 Bayesian GARCH Estimation

To describe the Bayesian framework, we first write

$$x_t = \epsilon_t \left(\omega_t \frac{v-2}{v} \sigma_t^2\right)^{1/2} \tag{6}$$

by following Geweke (1993). Let $w = (w_1, \ldots, w_T)'$, and $\alpha = (\alpha_0, \alpha_1)'$ and we regroup the unknown parameters as $\theta = (\alpha, \beta, v)'$. Upon defining the T × T diagonal matrix:

$$\Sigma = \Sigma(\theta, w) = \mathrm{diag}([w_t \frac{v-2}{v} \sigma_t^2(\alpha, \beta)] \tag{7}$$

the likelihood function of (θ, w), under the normal distribution, is defined as:

$$L(\theta, w \mid x) \propto [det(\Sigma)]^{-\frac{1}{2}} exp\left[-\frac{1}{2} x' \Sigma^{-1} x\right]$$

The parameters (θ, w) are random variables which are characterized by a prior density, denoted by p(θ, w). Inferences are made based on the posterior density defined by

$$p(\theta, w|x) = \frac{L(\theta, w|x)p(\theta, w)}{\int L(\theta, w|x)p(\theta, w)d\theta dw}. \tag{8}$$

After observing the data, the posterior distribution gives a probabilistic description of the knowledge about the model parameters.

Following Ardia (2010), we take the truncated normal prior distributions for the GARCH parameters α and β

$$p(\alpha) \propto \phi_2(\alpha|\mu_\alpha, \Sigma_\alpha) I[\alpha \in R_+^2] \; and \; p(\beta) \propto \phi_1(\beta|\mu_\beta, \Sigma_\beta) I[\beta \in R_+],$$

where ϕ_d is the d-dimensional normal density, μ. and Σ. are the hyper- parameters, and I[·] is the indicator function. Assuming that wt are inde- pendent and identically distributed as the inverse gamma with (v/2, v/2), the prior distribution of the vector w given v is

$$\mathrm{p(w|v)} = \left(\frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{\mathrm{v}}{2}\right)}\right)^{\mathrm{T}} \left(\prod_{t=1}^{T} w_t\right)^{-\frac{v}{2}-1} \exp(-\frac{1}{2}\sum_{t=1}^{T}\frac{v}{w_t})$$

The prior distribution of v is chosen as the translated exponential with λ > 0 and δ ≥ 2:

$$p(v) = \lambda exp[-\lambda(v - \delta)] I[v > \delta].$$

The mass of this prior is mostly concentrated near δ when λ is large and hence, the degree of freedom can be constrained in this manner. Deschamp (2006) points out that this prior density is useful in two ways. Bounding the degrees of freedom away from two may potentially be important from a numerical perspective to avoid a rapid divergence of the conditional variance. Next, the normality of the errors can be estimated while allowing the prior to remain reasonably constrained, which may allow for better convergence of the sampler.

Assuming the prior independence among the parameters, the joint prior distribution is then

$$p(\theta, w) = p(\alpha)p(\beta)p(w|v)p(v). \quad (9)$$

There is no closed form for the joint posterior distribution in (8) and no conjugate prior exists for this joint posterior density. Hence, we resort to the Markov chain Monte Carlo (MCMC) method for simulation to approximate the density of the posterior distribution. The MCMC sampling technique was initially introduced by Metropolis (1953) and was later generalized by Hastings (1970). The basic idea of this MCMC sampling method is based on the creation of a Markov chain $(\theta^{(0)}, w^{(0)}), \ldots, (\theta^{(k)}, w^{(k)})$ in the parameter space. Under some regularity conditions, as k goes to infinity, the asymptotic distribution of (θ(k), w(k)) will be (8). To implement the MCMC sampling technique, we used the Metropolis-Hastings (MH) algorithm. The details can be found in Chib (1995). This algorithm is used to update the GARCH parameters in blocks with one block for α and one block for β, while the parameter for degrees of freedom is sampled through an optimized rejection method from a translated exponential density defined earlier. This process is incorporated in the R package ***bayesGARCH*** for its MCMC sampler which uses the approach of Ardia (2008).

For $p(\alpha)$, we specify two cases for the variance-covariance matrix Σ_α:

$$\begin{bmatrix} 1000 & 0 \\ 0 & 1000 \end{bmatrix} \; and \; \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}$$

The first specification results in a diffuse (non-informative) prior density, whereas the second puts a heavy weight near 0 where the values of α will most likely to lie in $[0, 0.2] \times [0, 0.2]$. Similarly, the variance Σ_β for *p(β)* was set to be 1000 and 0.01. Both prior means μ_α and μ_β were set to 0.

2.3 Model Assessment

In the frequentist setting, we assumed $x_t = \sigma_t \epsilon_t$ with ϵ_t having the mean 0 and standard deviation 1. Therefore,

$$E(x_t^2) = E[E(x_t^2 \mid F_{t-1})] = E[E(\sigma_t^2 \epsilon_t^2 \mid F_{t-1})] = \sigma_t^2,$$

where, in practice, F_t denotes the past financial information up to time t. This is also true under the Bayesian setting because ϵ_t and w_t are independent, and $E(wt) = \frac{v}{v-2}$. Using this fact and the fact that the true squared volatility σ_t^2 is unknown when we deal with the actual datasets, we have used the squared series as a proxy for the squared volatility. Hence, we measure the mean square error (MSE) and the mean absolute deviance error (MADE) by

$$MSE = \frac{1}{T-p}\sum_{t=p+1}^{T}\left(\hat{\sigma}_t^{\,2} - x_t^2\right)^2 = \frac{1}{T-p}\sum_{t=p+1}^{T} a_t^2$$

and

$$MADE = \frac{1}{T-p}\sum_{t=p+1}^{T}|\hat{\sigma}_t^{\,2} - x_t^2| = \frac{1}{T-p}\sum_{t=p+1}^{T} a_t,$$

where $a_t = |\hat{\sigma}_t^{\,2} - x_t^2|$. As another measure of accuracy, we've used the directional accuracy (DA), which is defined by:

$$DA = \frac{1}{T-p}\sum_{t=p+1}^{T} d_t$$

$$d_t = \begin{cases} 1, & if\ (x_t^2 - x_{t-1}^2)(\hat{\sigma}_t^{\,2} - \hat{\sigma}_{t-1}^2) > 0 \\ 0, & otherwise \end{cases}$$

The DA gives the average direction of the forecast volatility by measuring the correctness of the turning point forecasts.

To test for significance in forecasting accuracy, we carried out the Diebold and Mariano (DM) test proposed by Diebold and Mariano (1995). The underlying hypotheses associated with this test are

$$H_0: E(z_t^{row}) = E\left(z_t^{column}\right)\ vs.\ H_1: E(z_t^{row}) \neq E\left(z_t^{column}\right),$$

where z_t^{row} and $z_t^{column}$are the squared deviance a_t^2 (and absolute deviance a_t) from the models in the row and the column, respectively. Hence, the null hypothesis indicates the "equal accuracy" between the two approaches. In large samples, the DM statistic

$$DM = \frac{\sqrt{T-p}(\bar{z} - \mu_z)}{2\pi f_z(0)} \sim N(0,1)$$

Where $\bar{z} = \frac{1}{T-p}\sum_{t=p+1}^{T} z_t = \frac{1}{T-p}\sum_{t=p+1}^{T}[z_t^{row} - z_t^{column}]$,$f_z(0) = \frac{1}{2\pi}\sum_{\tau=-\infty}^{\infty}\gamma_z(\tau)$ is the spectral density of the loss differential at frequency 0, and $\gamma_z(\tau) = E[(z_t - \mu)(z_{t-\tau} - \mu)]$ is the auto-covariance function at τ.

3. Results

In this section, we compare the predictive potentials of the GARCH(1,1) model under the frequentist and Bayesian methods using six daily exchange rates. We consider the daily exchange rates of six major currencies against US dollars. These currencies are Euro (EUR), Japanese yen (JPY), Pound sterling (GBP), Australian dollar (AUD), Swiss franc (CHF), and Canadian dollar (CAD). We analyze the most traded pairs of currencies, commonly called the *Majors*. The Majors are EUR/USD, GBP/USD, USD/JPY, AUD/USD, USD/CAD, and USD/CHF. Except for the EUR/USD pair,

Table 1: Numerical summary of the foreign exchange return series.

Exchange Rate	T	Min.	Median	Mean	Max.	Std. Dev.	Skewness	Kurtosis
EUR/USD	3635	-0.05	0.00	0.00	0.03	0.01	-0.12	2.06
GBP/USD	10656	-0.05	0.00	0.00	0.05	0.01	0.20	4.74
USD/JPY	10650	-0.06	0.00	0.00	0.10	0.01	0.70	9.79
AUD/USD	10649	-0.01	0.00	0.00	0.19	0.04	3.00	86.11
USD/CAD	10662	-0.04	0.00	0.00	0.05	0.01	0.09	12.45
USD/CHF	10656	-0.09	0.00	0.00	0.05	0.01	-0.13	5.44

all exchange rates start from January 4, 1971 and end at June 14, 2013. Since the Euro was introduced on January 1, 1999 in the financial market, the EUR/USD data set starts from January 4, 1999. These datasets can be downloaded from

http://research.stlouisfed.org/fred2/categories/158.

Several numerical summaries for the datasets are given in Table 1. It is noticeable that the skewness and kurtosis of AUD/USD are very high. This indicates that the distribution of the return series may be right-skewed and have fat tails. The fat-tail can also be noted from other datasets except for EUR/USD.

We've conducted some preliminary analyses of the datasets. Table 2 shows the results from Ljung-Box test based on squared return series. Except for EUR/USD when Q(1) = 2.103, the results indicate significant serial dependence.

Table 2: Ljung-Box Q Statistics based on square return series. P-values are in the parentheses.

	EUR/USD	GBP/USD	USD/JPY	AUD/USD	USD/CAD	USD/CHF
Q(1)	2.103	163.020	129.083	14.624	357.185	124.861
	(0.147)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Q(5)	53.035	1111.832	256.326	102.310	2406.758	470.215
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Q(10)	337.677	1955.381	331.681	148.410	5228.167	639.815
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Q(20)	689.314	3463.889	500.390	224.587	10978.330	959.078
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Q(50)	1557.367	6280.076	720.771	294.405	19985.733	1351.073
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)

The time series plots of the return series are shown in Figure 1. It can be seen that volatility clustering are present in the datasets. Also, the variability increases in the USD/CAD dataset. These may indicate that the datasets may not be stationary. Figure 2 shows the autocorrelation function (ACF) for the squared return series for each dataset. It is evident that the squared series seems to be serially correlated indicating a possible dependence at a higher moment.

For each dataset, the in-sample data consist of the first 70% of the dataset to fit the model and the out-of-sample data contain the last 30% to test the model. In practice, in-sample measures do not mean much since we are interested in predictive nature of the model. Table 3 gives the comparison based on these three measures for both in-sample and out-of- sample periods. Under the out-of-sample measures, the best measure that corresponds to the method is printed in bold for each dataset. The results indicate the frequentist approaches are generally better than the Bayesian approaches.

Table 4 gives the results based on DM test statistics. The values in each table are the test statistics indicating the significance of the model in the row versus the model in the column. In the table, we are primarily interested in the four upper-right entries for each dataset. A negative value indicates that the frequentist maximum likelihood approach gives smaller average error and hence, it is a better method. Most values are negative except for USD/CAD under MADE.

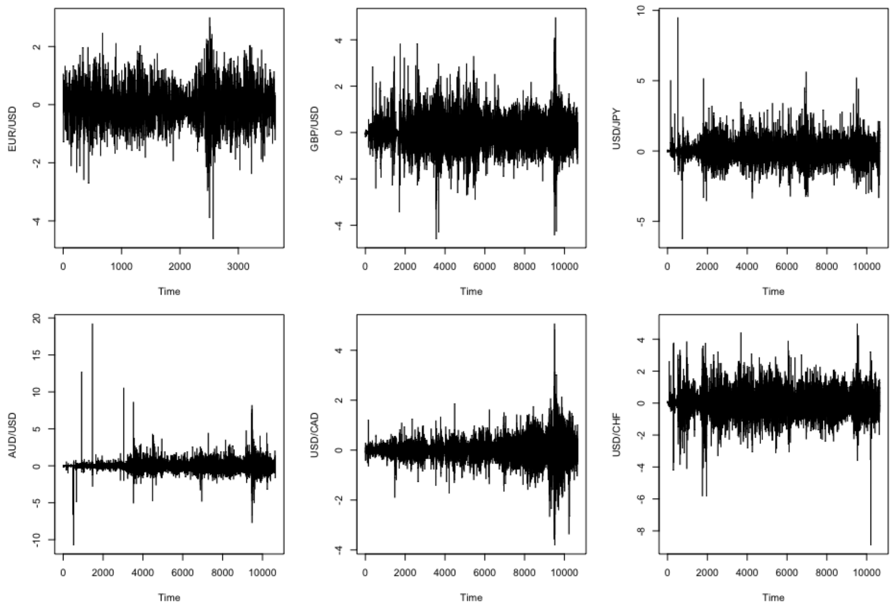

Figure 1: Time series plots for the return series in percent (%).

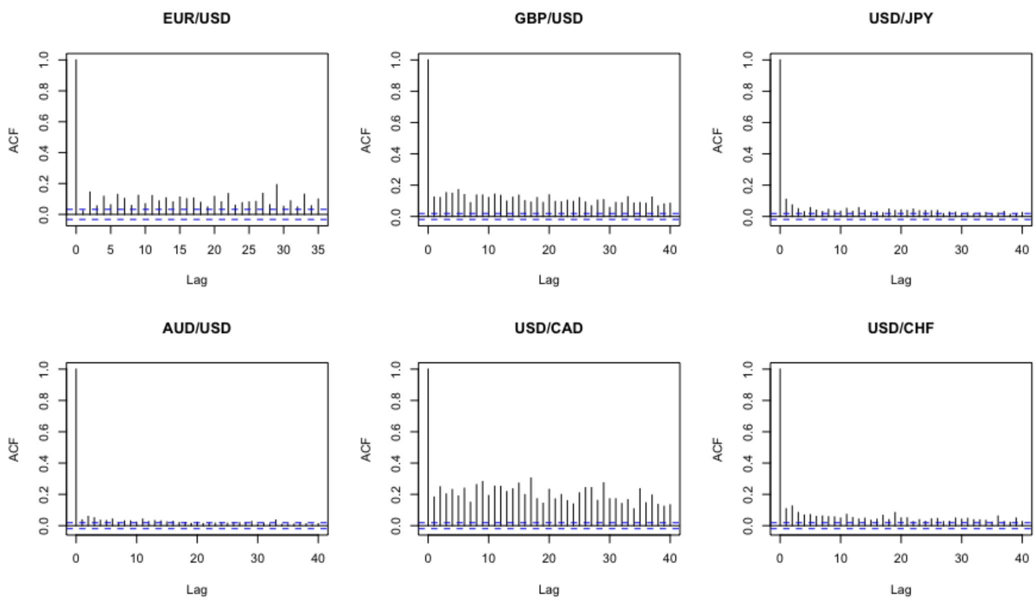

Figure 2: Autocorrelation function (ACF) plot for the squared return series.

Table 3: In-sample and out-of-sample measures for MSE, MADE and DA are shown. ML.normal and ML.t-dist denote the frequentist maximum likelihood method under normal and t-distribution likelihood functions, respectively. MH.diffuse and MH.t-normal denote the Bayesian method under diffuse and truncated normal prior distributions.

Dataset	Method	In-sample			Out-of-sample		
		MSE	MADE	DA	MSE	MADE	DA
EUR/USD	ML.normal	0.5838	0.4229	0.2718	**0.8724**	**0.4745**	0.2865
	ML.t-dist	0.5839	0.4230	0.4945	**0.8724**	**0.4745**	**0.2883**
	MH.diffuse	0.5847	0.4323	0.2722	0.8756	0.4772	0.2865
	MH.t-normal	0.5847	0.4319	0.2722	0.8755	0.4767	0.2865
GBP/USD	ML.normal	0.7560	0.4071	0.3104	**0.9234**	0.4023	**0.2782**
	ML.t-dist	0.7559	0.4051	0.4962	0.9243	**0.4009**	0.2776
	MH.diffuse	0.9523	0.5566	0.3045	1.1753	0.5480	0.2779
	MH.t-normal	0.9492	0.5546	0.3045	1.1717	0.5460	0.2779
USD/JPY	ML.normal	2.6038	0.5254	0.3111	**1.1365**	0.4905	**0.2806**
	ML.t-dist	2.6049	0.5212	0.4964	1.1400	**0.4881**	0.2769
	MH.diffuse	3.5084	0.8410	0.3121	1.6961	0.8013	0.2787
	MH.t-normal	3.5063	0.8403	0.3121	1.6946	0.8006	0.2787
AUD/USD	ML.normal	27.7762	0.4585	0.3006	**7.0678**	**0.7561**	**0.2697**
	ML.t-dist	30.1796	0.4792	0.5060	7.3948	0.8058	0.2794
	MH.diffuse	32.0382	0.6075	0.2848	7.5718	0.9552	0.2653
	MH.t-normal	32.0048	0.6050	0.2846	7.5497	0.9517	0.2653
USD/CAD	ML.normal	0.0257	0.0727	0.3021	0.9376	0.3975	0.2815
	ML.t-dist	0.0257	0.0727	0.5078	**0.9356**	0.3972	**0.2834**
	MH.diffuse	0.0255	0.0712	0.2949	1.0150	**0.3448**	0.2796
	MH.t-normal	0.0255	0.0712	0.2950	1.0155	**0.3448**	0.2799
USD/CHF	ML.normal	1.6863	0.6212	0.2972	2.9657	0.5738	**0.2845**
	ML.t-dist	1.6877	0.6192	0.4944	**2.9641**	**0.5717**	0.2851
	MH.diffuse	1.7808	0.6834	0.2990	3.0893	0.6293	0.2854
	MH.t-normal	1.7841	0.6849	0.2982	3.0952	0.6307	0.2851

Table 4: Diebold-Mariano (DM) test statistics for average loss differential based on MSE and MADE . Negative values signify better forecast in the corresponding row. And denote significance at 5% and 1%, respectively. ML.normal and ML.t-dist denote the frequentist maximum likelihood method under normal and t-distribution likelihood functions, respectively. MH.diffuse and MH.t-normal denote the Bayesian method under diffuse and truncated normal prior distributions.

MSE				
Dataset		ML.t-dist	MH.diffuse	MH.t-normal
EUR/USD	ML.normal	-0.442	-1.101	-1.131
	ML.t-dist	–	-1.098	-1.128
	MH.diffuse	–	–	1.191
GBP/USD	ML.normal	-1.878	-4.711**	-4.730**
	ML.t-dist	–	-4.688**	-4.708**
	MH.diffuse	–	–	5.651**
USD/JPY	ML.normal	-1.567	-8.614**	-8.621**
	ML.t-dist	–	-8.640**	-8.648**
	MH.diffuse	–	–	11.721**
AUD/USD	ML.normal	-0.82	-0.522	-0.541
	ML.t-dist	–	-0.214	-0.241
	MH.diffuse	–	–	2.174*
USD/CAD	ML.normal	1.473	-1.582	-1.576
	ML.t-dist	–	-1.612	-1.607
	MH.diffuse	–	–	-4.370**
USD/CHF	ML.normal	0.527	-3.437**	-3.417**
	ML.t-dist	–	-3.408**	-3.391**
	MH.diffuse	–	–	-3.146**

MADE				
		ML.t-dist	MH.diffuse	MH.t-normal
EUR/USD	ML.normal	-1.17	-1.338	-1.658
	ML.t-dist	–	-1.328	-1.649
	MH.diffuse	–	–	13.493**
GBP/USD	ML.normal	13.938**	-23.72**	-23.829**
	ML.t-dist	–	-23.88**	-23.987**
	MH.diffuse	–	–	27.237**
USD/JPY	ML.normal	5.456**	-34.950**	-34.977**
	ML.t-dist	–	-35.334**	-35.360**
	MH.diffuse	–	–	43.493**
AUD/USD	ML.normal	-4.177**	-9.339**	-9.426**
	ML.t-dist	–	-8.387**	-8.491**
	MH.diffuse	–	–	13.378**
USD/CAD	ML.normal	1.87	9.240**	9.248**
	ML.t-dist	–	9.181**	9.190**
	MH.diffuse	–	–	4.302**

USD/CHF	ML.normal	9.331**	-17.798**	-17.794**
	ML.t-dist	–	-18.370**	-18.391**
	MH.diffuse	–	–	-11.848**

4.Conclusion

Our main interest in this study was to compare the frequentist and Bayesian estimation approaches using the GARCH(1,1) as a basis model. In contrary to the existing literature, we have found that the frequentist method pro- vides better predictive potential than the Bayesian method. We considered six foreign exchange rate datasets. We computed MSE, MADE, and DA to compare different model outcomes and the out-of-samples indicate that the frequentist performed better. We also carried out DM test to observe the significance in these results. We have observed that it remains true, in general, that the frequentist provide more accurate predictive potential than the Bayesian approach. Finally, the current study is limited to the GARCH(1,1) as the basis model; however, one can use other basis model such as Exponential GARCH or Integrated GARCH models as well.

Acknowledgement

D. Yeung and S. S. Chung would like to thank Research Experience of Undergraduates (REU) NSF Grant #1460151 for supporting their research.

References

[1] Albers, C. J., Kiers, H. A., and van Ravenzwaaij, D. (2018). Credible Confidence: A pragmatic view on the frequentist vs Bayesian debate. Collabra: Psychology, 4(1).

[2] Billsus, D. and Pazzani M. J. (1998). Learning collaborative filters. In Proceedings of the 15th International Conference on Machine Learn- ing, 46-54. San Francisco, California.

[3] Bozios, T., Lekakos, G., Skoularidou, V. and Chorianopoulos K. (2001). Advanced techniques for personalized advertising in a digital TV en- vironment: the iMEDIA system. In Proceedings of The E-business and E-work Conference, 1025-1031. Venice, Italy.

[4] Burke, R. (2002). Hybrid recommender systems: survey and experiments. User Modeling and User-Adapted Interaction 12, 331-370.

[5] Ardia, D. (2008). Financial Risk Management with Bayesian Estimation of GARCH Models: Theory and Applications. In: Lecture Notes in Economics and Mathematical Systems, volume 612. Springer-Verlag, Berlin, Germany.

[6] Ardia, D. and Hoogerheide, L. F. (2010). Bayesian estimation of the GARCH(1,1) model with student-t innovations. The R Journal, 2(2):4147.

[7] Bauwens, L. and Lubrano, M. (1998). Bayesian inference on GARCH models using the Gibbs sampler. Simulation Methods in Economet- rics, 1(1):2346.

[8] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedas- ticity. Journal of Econometrics, 31(3):307327.

[9] Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. The American Statistician, 49(4):327335.

[10] Deschamps, P. J. (2006). A flexible prior distribution for Markov switch- ing autoregressions with student-t errors. Journal of Econometrics, 133(1):153190.

[11] Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. Journal of Business and Economic Statistics, 13(3):253263.

[12] Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates for the variance of united kingdom inflation. Econometrica, 50(4):9871008.

[13] Fan, J. and Yao, Q. (2005). Nonlinear Time Series. Springer.

[14] Geweke, J. F. (1993). Bayesian treatment of the independent student-t linear model. Journal of Applied Econometrics, 8(S1):1940.

[15] Ghalanos, A. (2011). Rsolnp: The General Non-linear Optimization.

[16] Ghalanos, A. (2016). rugarch: The Univariate GARCH Models.

[17] Hastings,W. K. (1970). Monte carlo sampling methods using Markov chains and their applications. Biometrika, 57(1):97109.

[18] Hoogerheide, L. F., Ardia, D., and Corre, N. (2012). Density prediction of stock index returns using garch models: Frequentist or Bayesian estimation? Economics Letters, 116:322325.

[19] Huang, S. F., Liu, Y. C., and Wu, J. Y. (2012). An empirical study on implied GARCH models. Journal of Data Science, 10(2012), 87-105.

[20] Jahufer, A. (2015). Choosing the best performing GARCH model for Sri Lank stock market by non-parametric specification test. Journal of Data Science, 13(2015), 457-472.

[21] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equations of state calculations by fast computing machines. Journal of Chemical Physics, 21(6):10871092.

[22] Nakatsuma, T. and Tsurumi, H. (1996). ARMA-GARCH models: Bayes estimation versus MLE, and Bayes non-stationarity test. Rutgers Uni- versity Working Paper, (9619).

[23] Nakatsuma, T. and Tsurumi, H. (1999). Bayesian estimation of ARMA- GARCH model of weekly foreign exchange rates. Asia-Pacific Finan- cial Markets, 6(1):7184.

[24] Samaniego, F. J. (2010). A comparison of the Bayesian and frequentist approaches to estimation. Springer Science & Business Media.

[25] Sigauke, C. (2016). Volatility modeling of the JSE all share index and risk estimation using the Bayesian and frequentist approaches. Economics, Management and Financial Markets, 11(4):3348.

[26] Tsay, R. S. (2010). Analysis of Financial Times Series. Wiley, 3rd edition.

[27] Vrontos, I. D., Dellaportas, P., and Politis, D. N. (2000). Full Bayesian inference for GARCH and EGARCH models. Journal of Business and Economic Statistics, 18(2):187198.

[28] Wagenmakers, E. J., Lee, M., Lodewyckx, T., and Iverson, G. J. (2008). Bayesian versus

frequentist inference. In Bayesian evaluation of infor- mative hypotheses (pp. 181-207). Springer, New York, NY.

[29] Wang, H., Yu, Y., and Li, M. (2010). On intraday Shanghai Stock Ex- change Index. Journal of Data Science, 8(2010), 413-427.

Appendix

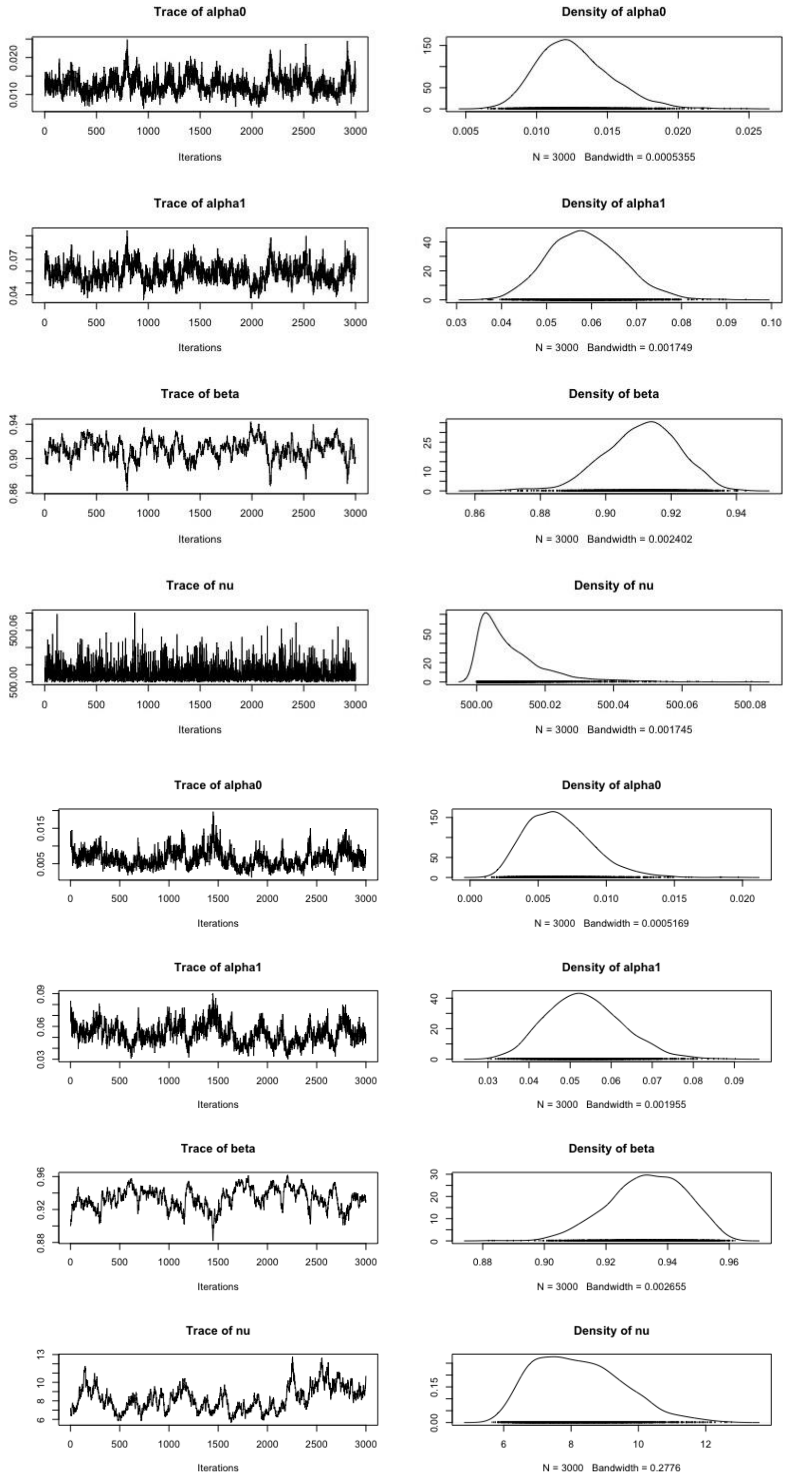

Figure 3: Trace plots and densities from diffuse prior (top 8) and truncated normal (bottom 8) for EUR/USD dataset.

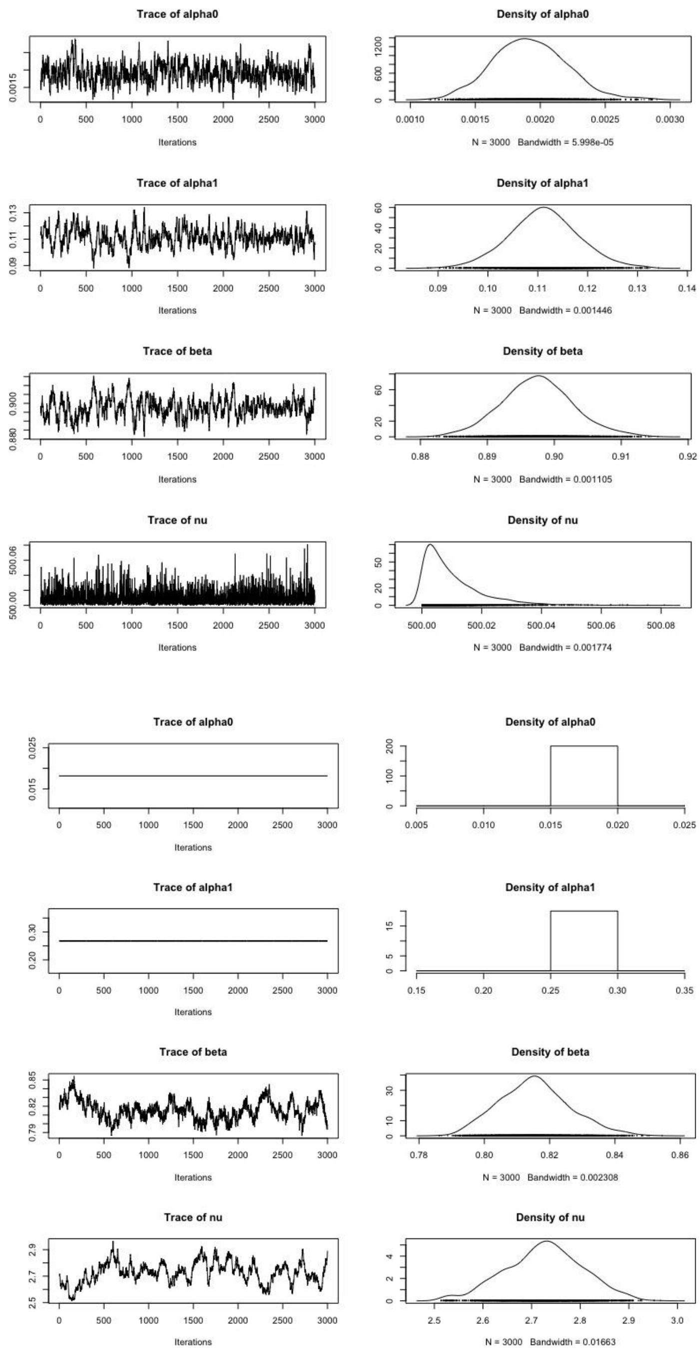

Figure 4: Trace plots and densities from diffuse prior (top 8) and truncated normal (bottom 8) for GBP/USD dataset.

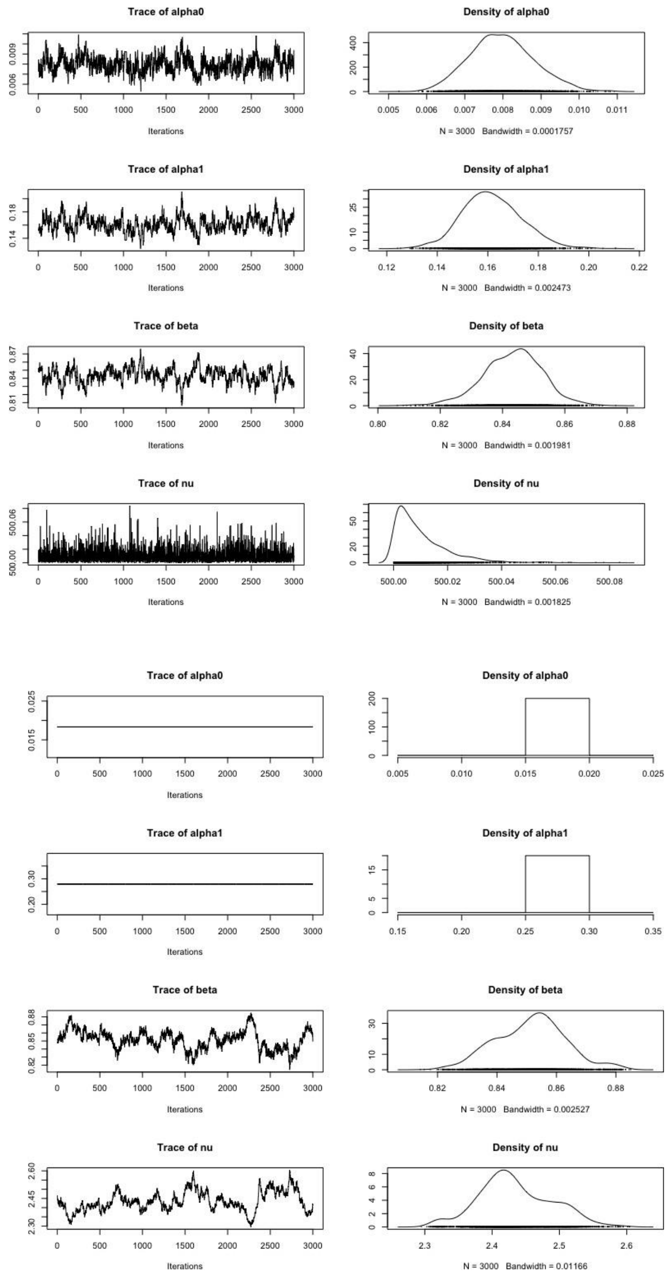

Figure 5: Trace plots and densities from diffuse prior (top 8) and truncated normal (bottom 8) for USD/JPY dataset.

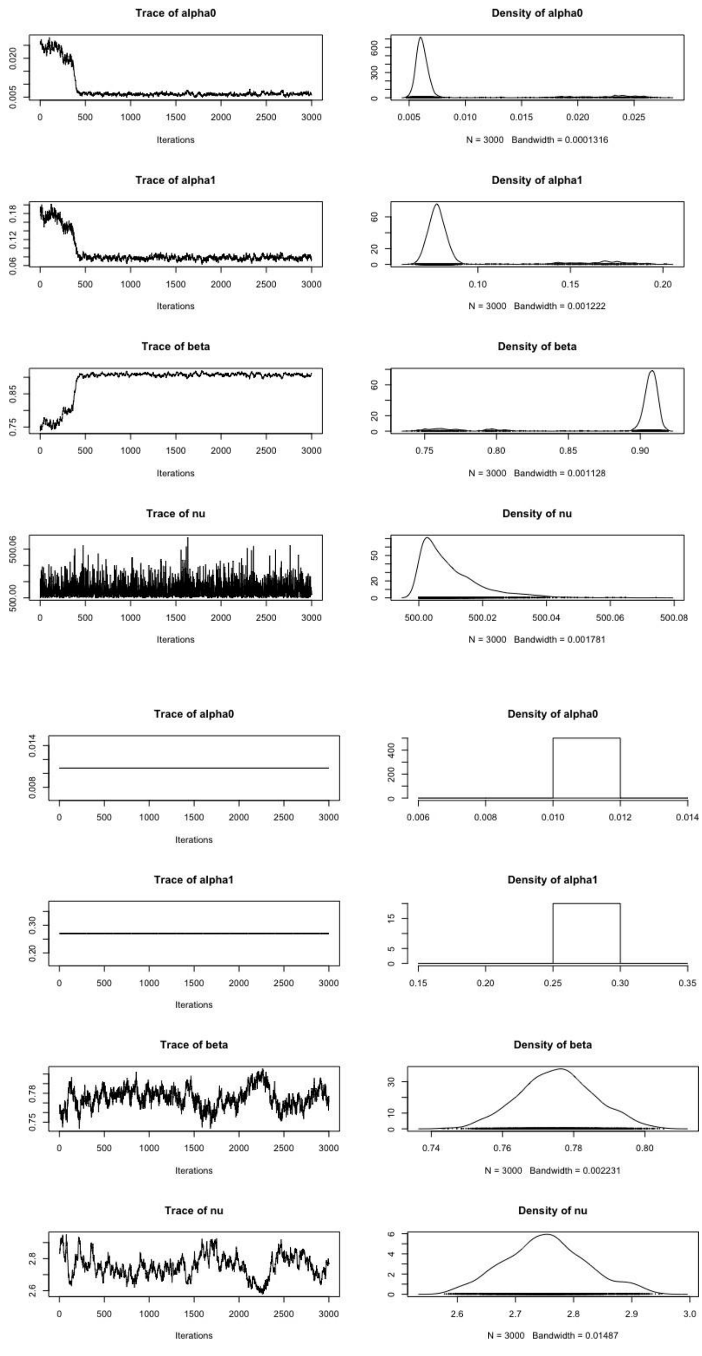

Figure 6: Trace plots and densities from diffuse prior (top 8) and truncated normal (bottom 8) for AUD/USD dataset.

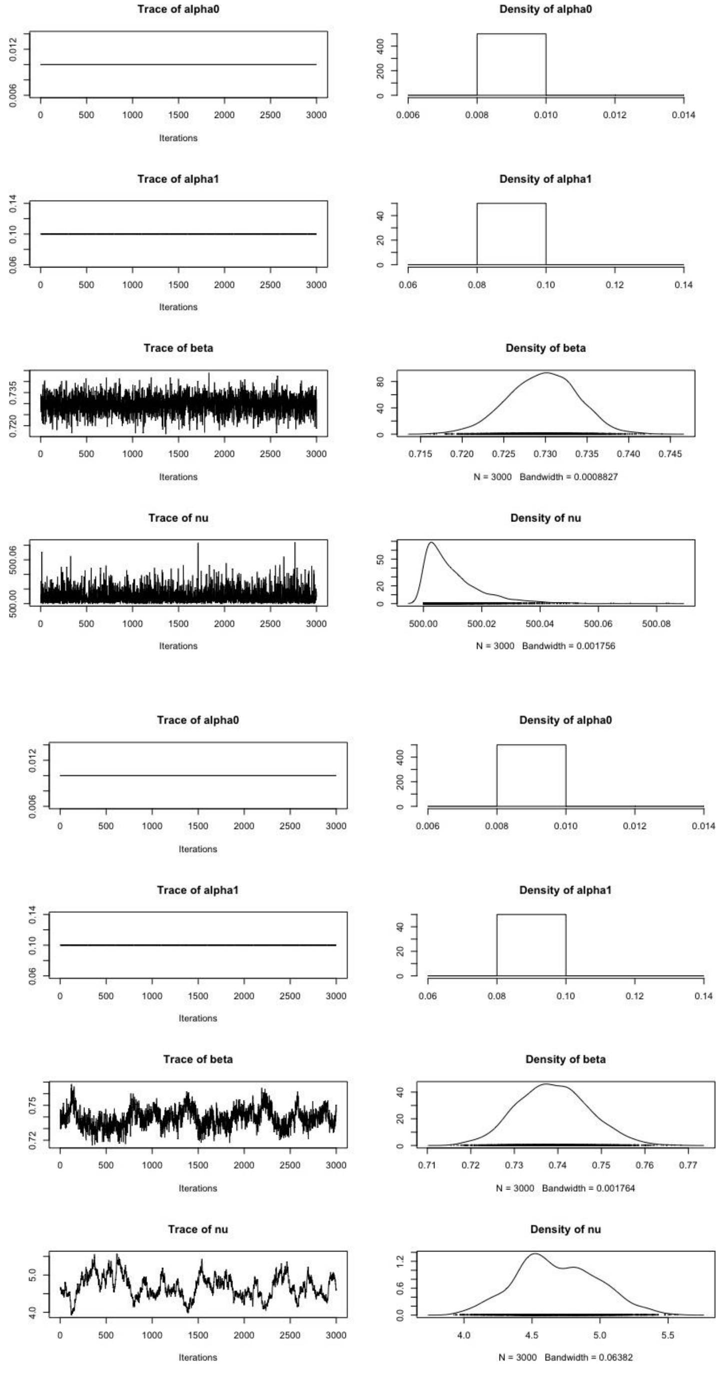

Figure 7: Trace plots and densities from diffuse prior (top 8) and truncated normal (bottom 8) for USD/CAD dataset.

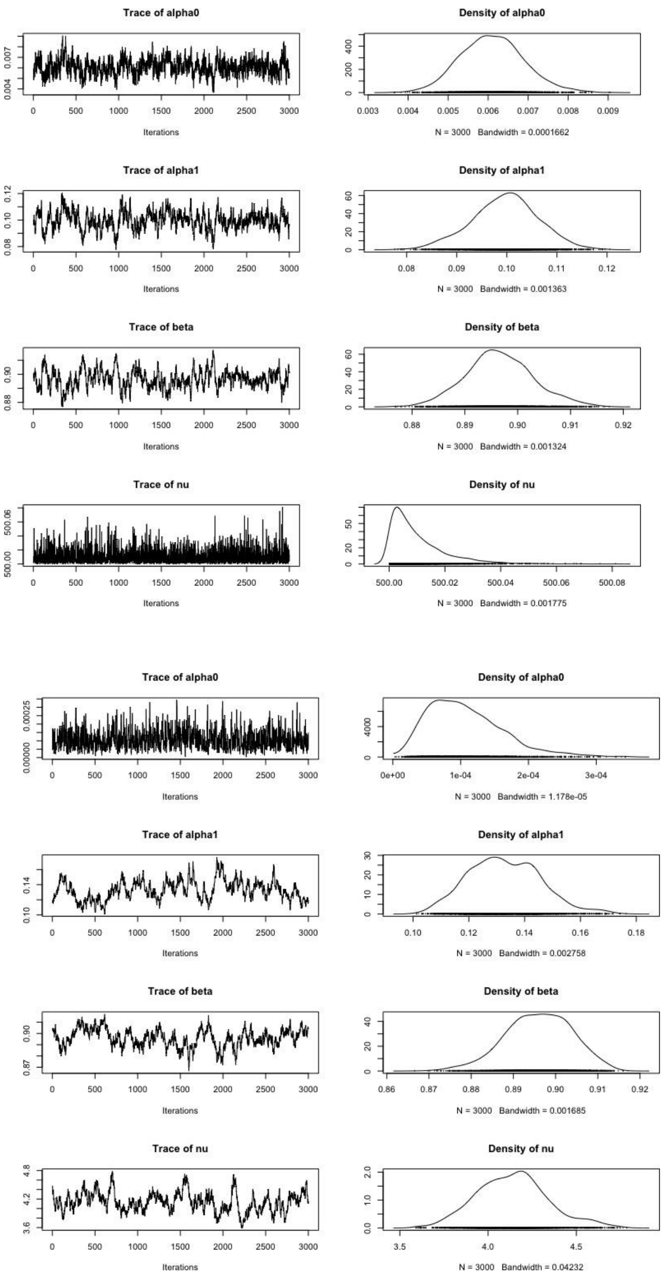

Figure 8: Trace plots and densities from diffuse prior (top 8) and truncate normal (bottom 8) for USD/CHF dataset.

Steve Chung (Corresponding author)
Department of Mathematics
California State University, Fresno
Fresno, CA 93740, USA
Email: schung@csufresno.edu
Phone: 559-278-2462
Fax: 559-278-2872

Jalen Harris and Christopher Newmark
Department of Mathematics
California State University, Fresno
Fresno, CA 93740, USA
Email: jalenharris@mail.fresnostate.edu; cnewmark@mail.fresnostate.edu

Diana Yeung
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556, USA
Email: Diana.Y.Yeung.7@nd.edu