Journal of the Nigerian Mathematical Society
Vol. 39, Issue 2, pp. 183-209, 2020

BAYESIAN ESTIMATION OF TIME-VARYING PARAMETERS IN THE PRESENCE OF DISCOUNTED EVOLUTION VARIANCE

O. O. AWE[1] AND A. A. ADEPOJU

ABSTRACT. Considerable attention has been devoted in literature to the estimation of linear models with constant location parameters. However, many phenomena in real life situations exhibit a non-linear time-varying pattern, indicating a need to adopt a Bayesian dynamic model and deal with the complexity involved in estimating the resulting time-varying parameters. In this paper, we present a novel application involving the estimation of time-varying parameters in dynamic state space models in the presence of discounted evolution variance. A conceptual derivation of the posterior distribution of time-varying parameters was done, with the application of a proposed discounting technique examined with simulated and crude oil exportation data. The results showed substantial time-variation in the slope parameters associated with the studied location parameters, thereby highlighting the empirical relevance and advantage of the discounting method as well as its computationally less intensive approach.

Keywords and phrases: State space model, discount factors, time-varying parameters, evolution variance, dynamic model.

2010 Mathematical Subject Classification: 62M10; 62P20

1. Preliminary

Dynamic state space models have gained tremendous popularity and application in many fields including Space Science, Physics, Mathematical Economics, Time Series Analysis and Biostatistics. They are suitable for modeling a wide array of data (univariate and multivariate) in the presence of non-stationarity, structural changes and irregular patterns [1,2]. A model is generally said to be dynamic every time its variables (or parameters) are indexed by time or appear with different time lags [3]. Dynamic time series models

Received by the editors May 28, 2020; Revised July 28, 2020 ; Accepted: August 18, 2020

www.nigerianmathematicalsociety.org; Journal available online at https://ojs.ictp.it/jnms/

[1]Corresponding author

have been described in different ways in literature. Authors like [4,5] presented autoregressive models where the dynamic structure appears on the endogenous variable as

$$y_t = \alpha y_{t-1} + e_t \tag{1}$$

where y_t is the observed time series, α is the slope parameter, y_{t-1} is the lagged value of the observed time series, and e_t is the stochastic error term. If the dynamic structure appears on the exogenous variable, it is known to be a distributed lag model [6] given as

$$y_t = \alpha + \theta_0 x_t + \theta_1 x_{t-1} + \theta_2 x_{t-2} + ... + \theta_n x_{t-n} + e_t \tag{2}$$

where y_t is the value at period t of the dependent variable y. α is the intercept term and θ_i is the lag weight to be estimated. Other forms of autoregressive distributed lag models where the dynamic structure appears on either the endogenous or exogenous variables were estimated in [5]. The dynamic structure can also appear on the error process of the model (e.g MA and ARMA models of [7] given as

$$y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \tag{3}$$

where μ is the mean of the series and ε are white noise error terms.

The first Bayesian approach to time series analysis and forecasting stem from [8] and is based on the dynamic linear model (also known as state space models) [9] in which the dynamic structure affects the location parameters in the model. Introducing time-varying parameters into dynamic models can lead to different levels of complexities. For instance, many phenomena in macro econometrics exhibit a time-varying pattern, indicating a need to reduce the resulting complexity in the estimation of the evolution variance through appropriate estimation algorithms and proper choice of discount values (λ) which many previous studies have failed to effectively address. This is the main focus of this present work.

Analysis of data that vary over time (or space) poses a great deal of challenge to Data Scientists and Econometricians. In recent times, estimation of time-varying parameters in econometric models has become more relevant especially as the length of the observed time series increases and the series itself is subject to changes in the dynamic structure. Particular examples can be found in world economic time series where key quarterly or monthly indicators are commonly available from the 1950s and cover periods of different economic situations. For example, since the 1950s, there have been periods of strong economic growth in the 1950s and 1960s, periods with oil crises in the 1970s, periods of major monetary policy changes in the 1980s, rapid changes of financial markets in the 1990s and the collapse of the financial and banking systems around the years 2008 and 2009 [10, 11]. Although not all economic structures are subject to change due to some erratic developments [13], it is expected that the dynamic properties of longer time series require parameters that are allowed to change over time. [9].

The works of researchers like [1], [12] and [2] provides a Bayesian alternative to the classical approaches to modeling time series data by Bayesian state space models. For instance, it has been argued severally in literature that the parameters in econometric models cannot, in general, be expected to remain constant and hence models with time-varying parameters should be considered in almost all circumstances [13, 14]. The difficulty in estimating such models is however often exacerbated by the fact that the Statistician (or Econometrician) would have only some ideas regarding the most likely value that a parameter may assume [11]. Some of the rationales behind the ideals of such models are fully documented in [15] and [1]. One distinguishing novelty of this present work is the use of discount factors for estimating the volatility of the state evolution variance in dynamic state space time series models. This study would be a knowledge base to data scientists, modelers and researchers in the area of Bayesian dynamic model estimation. Our attempt to make use of the recursive forward filtering backward sampling technique within the Kalman filter framework in the presence of discounted evolution variance enhances a fast and efficient Markov chain Monte Carlo (MCMC) sampling of the time-varying state parameters which is applicable to many fields.

Essentially, this paper presents the mathematical formulation of the dynamic state space model for time series analysis, its implementation, computational guidelines, tuning of the discounting parameters as well as its practical forecasting applications in helping the research community to better understand its use and importance in estimating the evolution variance of the state space model. We then applied the model to the time series modeling of crude oil export in Nigeria. In the long run (after the COVID-19 pandemic), drop in crude oil prices may impact the economic stability and sustainability of many countries, especially those depending on crude oil exports like Nigeria. Previous analysis of crude oil export data in Nigeria were done with classical parametric Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) or fractional integration models [16–19]. The weakness of these models lie in the fact that they can not estimate the necessary time-varying parameters which were estimated in the present study. After this non-exhaustive introductory section, the rest of this paper is structured as follows: Section 2 is on the model specification and Bayesian methodology involved in estimating time-varying parameters. Section 3 deals with the empirical analyses and applications of the methods in this study. Section 4 is on discussion of results and some empirical insights, while Section 5 concludes the study.

2. Model Specification and Methodology

The proposed dynamic (state space) model specification takes the following form:

$$y_t = X_t\theta_t + v_t \qquad v_t \sim N(0, V), \tag{4}$$

$$\theta_t = G_t\theta_{t-1} + w_t \qquad w_t \sim N(0, W_t), \tag{5}$$

$$\theta_0 \sim N(m_0, C_0).$$

Equation (4) is known as the observation equation while equation (5) is the evolution equation. G_t is a known matrix of order $p \times p$ that determines how the observation and state equations evolve in time . We assume that all v_t's are independent from the w_t's. Since each parameter at time t only depends on results from time $t-1$, the state parameters are time-varying and constitute a Markov chain. X_t is a matrix of observed explanatory variables of a known order. It is assumed that information decays arithmetically through the addition of future evolution error variance which is estimated via discounting.

Parameters of interest to be estimated in the model are the time-varying parameter θ_t, the error variances V and W_t, and the one-step-ahead forecasts error f_t. V is assumed to be distributed inverse-gamma apriori and is estimated using the Bayesian framework discussed in the next section, while we estimate W_t using the proposed discounting method. In contrast to the Box-Jenkins methodology, which still plays an important role in time series analysis [7], this specified Bayesian dynamic model approach allows for structural analysis of univariate as well as multivariate problems. The different components of time series, such as the trend and seasonal terms can be modeled explicitly. They do not have to be removed prior to the main analysis as is the case in the Box-Jenkins methodology.

2.1. **Estimating Time-Varying Parameters: Recursive Forward Filtering Backward Sampling (RFFBS) Approach.**
In this section, we consider a procedure for estimating dynamic state space models. This approach makes use of the RFFBS algorithm within the Kalman filter framework to improve the efficiency of the Gibbs sampling. The main idea of this procedure is to make use of the Markov's property of the specified model and the structure of the state transition equation so that

$$P(S_t|S_{t+1}, D_n) = P(S_7|S_6, D_t, \nu_{t+1}, ..., \nu_n) \tag{6}$$

where S_t denotes the state variable at time t and ν_j is the 1-step-ahead prediction error. This recursive method allows us to draw the state vectors jointly. Due to the Markovian structure of the time-varying parameter θ_t, it is estimated by computing the predictive and filtering distributions of θ_t recursively starting from the prior $\theta_0 \sim N(m_0, C_0)$. Consider a vector of unknown regression slope parameters $\theta_t = (\theta_1, ..., \theta_p)$, the Gibbs sampling algorithm employed proceeds by sampling recursively the conditional posterior distribution where the most recent values of the conditioning parameters are used. Assume that the observed response is represented by $y = (y_1, y_2, ..., y_T)$ where T denotes the size of the series, the model is estimated by simulating the distribution of the parameters of interest, given the data.

Following the Bayesian paradigm, the specification of the model is complete only after specifying the prior distribution of all the unknown quantities of interest in the model [20]. We assign a distribution to θ_t at time t=0, conditional on all the information available before any observation is made. Let D_0 be the set containing all this information, then the prior distribution is $\theta_0|D_0 \sim N(m_0, C_0)$ where m_0 and C_0 are known vector and matrix respectively. Next, an update is made for θ_1 and D_0 which is also normally distributed. Based on this update, the one step-ahead forecast follows from the conditional distribution $y|\theta_0, D_0$. Once the value of y_1 at time $t = 1$ is known, the posterior distribution of θ_1 is obtained recognizing that the information available at time $t = 1$ is $D_1 = y_1, D_0$. The inference is made in this recursive fashion for every time t. The Kalman filter was used to calculate the mean and variance of the

unobserved state θ_t, given the observations y_t. It is a recursive algorithm i.e the current best estimate is updated whenever a new observation is obtained. The filter prediction and update algorithm requires a few basic calculations of which only the conditional means and variances of the filtering and prediction density is stored in each step of the iteration [**?**, 21].
To describe the filtering procedure, let

$$m_t = E(\theta_t | D_t) \tag{7}$$

be the optimal estimator of θ_t based on D_t and let

$$C_t = E((\theta_t - m_t)(\theta_t - m_t)^T | D_t) \tag{8}$$

be the mean square error matrix of m_t. Let $\theta_{t-1}|y_{1:t-1} \sim N(m_{t-1}, C_{t-1})$, where $y_{1:t-1}$ denote all observations up to time $t-1$. Then the one-step-ahead predictive density $\theta_t|y_{1:t-1}$ is Gaussian with parameters:

$$E(\theta_t | y_{1:t-1}) = m_{t-1} \equiv A_t \tag{9}$$

$$Var(\theta_t | y_{1:t-1}) = C_{t-1} + W_t \equiv R_t \tag{10}$$

The one-step-ahead predictive density of $y_t|y_{1:t-1}$ is Gaussian with parameters:

$$f_t = E(y_t | y_{1:t-1}) = X_t A_t \tag{11}$$

$$Q_t = Var(y_t | y_{1:t-1}) = X_t R_t X_t' + V \tag{12}$$

The filtering density of θ_t given $y_{1:t}$ is Gaussian with parameters:

$$m_t = E(\theta_t | y_{1:t}) = A_t + R_t X_t' Q_t^{-1} e_t \tag{13}$$

$$C_t = Var(\theta_t | y_{1:t}) = R_t - R_t X_t' Q_t^{-1} X_t R_t' \tag{14}$$

where $e_t = y_t - f_t$ is the forecast error.

2.2. **RFFBS Algorithm and Gibbs Sampler.** Let $\theta_T = [\theta_0, \theta_1, ..., \theta_t]$, θ_t was estimated from the conditional density $p(\theta_t|y_T)$ which is denoted by

$$p(\theta_t, y_T) = p(y_T|\theta_T)p(\theta_T)$$

where $p(y_T|\theta_T)$ and $p(\theta_T)$ are given by

$$p(y_T|\theta_T) = \prod_{t=1}^{T} p(y_t|\theta_t)$$

and

$$p(\theta_T) = p(\theta_0)\prod_{t=1}^{T}(\theta_t|\theta_{t-1})$$

. $p(y_t|\theta_t)$ and $p(\theta_t|\theta_{t-1})$ were derived from the measurement and evolution equations (5) specified above to give

$$p(y_t|\theta_t) = (2\pi V)^{-\frac{1}{2}} exp\Big(-\frac{1}{2V}(y_t - x_t\theta_t)^2\Big)$$

where we

$$p(\theta_t|\theta_{t-1}) = (2\pi)^{-\frac{k}{2}}|W_t|^{-\frac{1}{2}} exp\Big(-\frac{1}{2}(\theta_t - G_t\theta_{t-1})'W_t^{-1}(\theta_t - G_t\theta_{t-1})\Big)$$

.

Since the probability distribution of update is proportional to the product of the time series measurement likelihood and the predicted state,

$$p(\theta_t|y_{1:t}) = \frac{p(y_t|\theta_t)p(\theta_t|\theta_{t-1}, y_{1:t-1})}{p(y_t|y_{1:t-1})} \qquad (15)$$

$$\propto p(y_t|\theta_t)p(\theta_t|\theta_{t-1}, y_{1:t-1})$$

The denominator, $p(y_t|y_{1:t-1})$ is constant relative to θ_t and thereby ignored. The posterior distribution was used to update the prior recursively until convergence is achieved. The new Recursive Forward Filtering Backward Sampling (RFFBS) algorithm adopted in this work is to allow for the implementation of a fast MCMC approach to the specified dynamic linear model. The forward filtering step is the standard Kalman filtering analysis which gives $p(\theta_t|D_t)$ at each t, for $t = 1, ..., n$. The backward sampling step uses the Markov's property specified in equation (6) above to sample θ_n^* from $p(\theta_n|D_n)$ and then for $t = 1, ..., n-1$, sample θ_t^* from $p(\theta_t|D_t, \theta_{t+1}^*)$ in order to generate samples from the posterior parameter structure. In particular, we denote

$$p(\theta_0, ..., \theta_T|D_T) = \prod_{t=0}^{T} p(\theta_t|\theta_{t+1}, ..., \theta_T, D_T)$$

and note that, by the Markov's property,

$$p(\theta_t|\theta_{t+1}, ..., \theta_T, D_T) = p(\theta_t|\theta_{t+1}, D_T)$$

then, the RFFBS algorithm proceeds as follows:

(1) Sample from $p(\theta_T|D_T)$ using the filtering density above. This distribution is $N(h_t, H_t)$ where:

$$h_t = m_t + C_t G_t' R_{t+1}^{-1}(\theta_{t+1} - a_{t+1}) \tag{16}$$

$$H_t = C_t - C_t G_t' R_{t+1}^{-1} G_t C_t' \tag{17}$$

(2) Sample from $p(\theta_{T-1}|\theta_T, D_T)$.
(3) Given $(\theta_t|D_t)$, obtain $W_t = C_t(1-\lambda)/\lambda$.
(4) Proceed recursively in this manner for $t+1, t+2$, and so on .
(5) Sample from $p(\theta_0, ..., \theta_T|D_T)$.

Since we sampled from $t = T$ to $t = 0$, recursively, this procedure is referred to as recursive backward sampling.

2.3. **Posterior Estimation of Unknown Observation Variance (V) with Independent Priors.** In the simulation excercise for estimating the static observational variance, V, the following Gibbs sampler of [10] was adopted with slight modifications:

(1) Choose an arbitrary starting point $\pi^{(0)} = (\pi_1^{(0)}, ..., \pi_p^{(0)})$ and set $i = 0$.
(2) Given $\pi^{(i)} = (\pi_1^{(i)}, ..., \pi_p^{(i)})$,
(a) generate $\pi_1^{(i+1)}$ from the conditional posterior distribution $\pi(\pi_1^{(i+1)}|\pi_2^{(i)}, ..., \pi_p^{(i)})$,
(b)generate $\pi_2^{(i+1)}$ from $\pi(\pi_2^{(i+1)}|\pi_1^{(i+1)}, \pi_3^{(i)}, ..., \pi_p^{(i)})$
(c)generate $\pi_3^{(i+1)}$ from $\pi(\pi_3^{(i+1)}|\pi_1^{(i+1)}, \pi_2^{(i+1)}, \pi_4^{(i)}, ..., \pi_p^{(i)}$,
(d)generate $\pi_4^{(i+1)}, ..., \pi_p^{(i+1)}$ in the same way as before.
(3) Set $i = i + 1$ and go to (2).

The Gibbs Sampling (GS) algorithm is an example of MCMC method described earlier. What is done in the GS algorithm is to break the joint posterior into conditional posteriors for which the analytical form of its density is known, we then sample sequentially and repeatedly from these conditionals. After a number of draws, we expect the joint sequence of conditional draws to converge to the desired joint posterior densities for all parameters. Each sequence of draws is then interpreted as the marginal posterior for a given parameter. The procedures and implementations of the Gibbs sampling algorithm are fully documented in [20] and [1]

.

Consider the observational equation specified in (4) above

$$y_t = X_t\theta_t + v_t, v_t \sim N(0, V), \tag{18}$$

and assume a normal prior for the parameter θ and inverse gamma prior for the parameter V, to sample from $V|\theta$ we impose a gamma prior on V^{-1} and derive the posterior hyperparameters. Let $V^{-1} \sim Gamma(a_0, b_0)$, then

$$V^{-1}|\theta \sim Gamma(a_0 + \frac{T}{2}, b_0 + \frac{1}{2}\sum_{t=1}^{T}(y_t - X_t\theta_t)^2)$$

Starting with

$$p(y|\theta, X) = (2\pi)^{-\frac{n}{2}} exp(-\frac{1}{2V}(y - X\theta)'(y - X\theta)) \tag{19}$$

the priors are given as follows:

$$\theta \sim N(\mu_0, V_0) \tag{20}$$

, and

$$V \sim IG(\nu_0, \tau_0) \tag{21}$$

μ_0 is the prior mean for θ and V_0 is the prior variance- covariance matrix for θ with

$$E(V) = \frac{\tau_0}{\nu_0 - 1} \tag{22}$$

,

$$V(V) = \frac{\tau_0^2}{(\nu_0 - 1)^2(\nu_0 - 2)} \tag{23}$$

We chose the form given in [22] where ν_0 and τ_0 are the shape and scale parameters respectively. Using Bayes rule to combine the priors (22) and (23) above with the likelihood via conjugate analysis and dropping all unrelated terms to the parameters of interest yields the following posterior kernels:

$$p(\theta, V|y, X) \propto (V)^{\frac{-n-2\nu_0-2}{2}} exp(-\frac{1}{2V}(2\tau_0))$$
$$\times\, exp\left(-\frac{1}{2}(\frac{1}{V}(y - X\theta)'(y - X\theta) + (\theta - \mu_0)'V_0^{-1}(\theta - \mu_0))\right) \tag{24}$$

First, obtain the posterior density of θ, conditional on V while treating σ^2 as a constant.
This leaves us with the posterior kernel:

$$p(\theta|V, y, X) \propto$$
$$exp(-\frac{1}{2}(\frac{1}{V}(y - X\theta)^{'}(y - X\theta) + (\theta - \mu_0)^{'}(V_0)^{-1}(\theta - \mu_0))). \tag{25}$$

Let

$$V_1 = (V_0^{-1} + \frac{1}{V}X'X)^{-1}$$

and

$$\mu_1 = V_1(V_0^{-1}\mu_0 + \frac{1}{V}X'Xb) = V_1(V_0^{-1}\mu_0 + \frac{1}{V}X'y)$$

Then from (25),

$$\begin{aligned}
&\frac{1}{\sigma^2}(y - X\theta)'(y - X\theta) + (\theta - \mu_0)^{'}V_0^{-1}(\theta - \mu_0)\\
= &\ \frac{1}{V}y'y + \theta'\frac{1}{V}X'X\theta - \frac{1}{V}y'X\theta - \theta'\frac{1}{V}X'y + \theta'V_0^{-1}\theta\\
- &\ \mu_0^{'}V_0^{-1}\theta - \theta^{'}V_0^{-1}\mu_0 + \mu_0^{'}V_0^{-1}\mu_0\\
= &\ \theta'(V_0^{-1} + \frac{1}{V}X^{'}X)\theta - \theta^{'}(V_0^{-1}\mu_0 + \frac{1}{V}X^{'}y)\\
- &\ (\mu_0^{'}V_0^{-1} + \frac{1}{V}y'X)\theta + \frac{1}{V}y^{'}y + \mu_0^{'}V_0^{-1}\mu_0\\
= &\ \theta^{'}V_1^{-1}\theta - \theta^{'}V_1^{-1}\mu_1 - \mu_1^{'}V_1^{-1}\theta\\
+ &\ \mu_1^{'}V_1^{-1}\mu_1 - \mu_1^{'}V_1^{-1}\mu_1 + \frac{1}{V}y'y + \mu_0^{'}V_0^{-1}\mu_0\\
= &\ (\theta - \mu_1)'V_1^{-1}(\theta - \mu_1) - \mu_1'V_1^{-1}\mu_1V_1^{-1}\mu_1 + \frac{1}{V}y'y\\
+ &\ \mu_0'V_0^{-1}\mu_0.
\end{aligned}$$

Therefore, the conditional posterior kernel in (25) can be written as :

$$p(\theta|V, y, X) \propto$$

$$exp(-\frac{1}{2}(\theta - \mu_1)'V_1^{-1}(\theta - \mu_1)exp(-\frac{1}{2}(\frac{1}{V}y'y + \mu_0'V_0^{-1}\mu_0 - \mu_1'V_1^{-1}\mu_1)) \tag{26}$$

Since none of the terms in the second exponent include θ, we simplify the full conditional distribution in (26) to

$$p(\theta|V, y, X) \propto exp(-\frac{1}{2}(\theta - \mu_1)'V_1^{-1}(\theta - \mu_1)) \tag{27}$$

Therefore, we have again,the kernel of a multivariate normal density , and we can say that

$$\theta|V, y, X \sim N(\mu_1, V_1)$$

where

$$V_1 = (V_0^{-1} + \frac{1}{V}X'X)^{-1}$$

and

$$\mu_1 = V_1(V_0^{-1}\mu_0 + \frac{1}{V}X'y)$$

to sample from.

Posterior Inference on V

In order to derive the conditional posterior density for V, we return to our original expression for the joint posterior given in (25). Ignoring terms that are not related to V, we have :

$$p(V|\theta, y, X) \propto (V)^{\frac{-n-2\nu_0-2}{2}} exp(-\frac{1}{2V}(2\tau_0 + (y - X\theta)'(y - X\theta))) \quad (28)$$

Comparing this expression with the kernel of the inverse gamma prior specified in (21), we have the kernel of another inverse gamma density: Hence

$$V|\theta, y, X \sim IG(\nu_1, \tau_1) \quad (29)$$

where

$$\nu_1 = \frac{2\nu_0 + n}{2}$$

and

$$\tau_1 = \frac{2\tau_0 + (y - X\theta)'(y - X\theta)}{2}$$

2.4. **Estimation of Evolution Variance (W_t) with Discount Values.** Consider the evolution equation in (5) above,

$$\theta_t = G_t\theta_{t-1} + w_t, w_t \sim N(0, W_t) \quad (30)$$

where W_t is the evolution variance and other parameters are as defined earlier. Let

$$\begin{aligned} V(\theta_{t-1}|D_{t-1}) &= V(G_t\theta_{t-1}|D_{t-1}) \\ &= G_tC_{t-1}G_t^T \\ &= C_{t-1} \end{aligned}$$

so that

$$Var(\theta_t|D_{t-1}) = C_{t-1} + W_t$$

The prior distribution for θ_{t-1} is

$$\theta_{t-1}|D_{t-1} \sim N(m_{t-1}, C_{t-1})$$

where $D_{t-1} = (y_1, y_2, ..., y_{t-1})$ and the prior distribution for θ_t is

$$\theta_t|D_{t-1} \sim N(m_{t-1}, Q_t)$$

where

$$Q_t = C_{t-1} + W_t$$

Therefore,

$$W_t = Q_t - C_{t-1} \tag{31}$$

We introduce the discount factor as a quantity λ such that

$$Q_t = C_{t-1}/\lambda \tag{32}$$

can be interpreted as the percentage of information that passes from time $t-1$ to t.

The idea of discounting is adopted in order give a natural interpretation to, and represent W_t as a proportion of the filtering distribution variance C_t while allowing it to vary through time and model changes in volatility [9, 21]. If C_t is large then there is high uncertainty in moving from θ_{t-1} to θ_t. Since W_t represents this uncertainty, it is necessary to model it as proportional to the filtering variance C_t. Therefore, we select the appropriate discount values from the grid $\lambda \in [0.01, 0.99]$ using the algorithm of [23]. As far as one-step-ahead forecasts are concerned, there is no need to refer to W_t explicitly in the model estimation since $Q_t = P_t/\lambda$ where $(P_t = C_{t-1})$. We examined multiple granularities of λ in the Gibbs sampler to see which one worked best. We then used the best discount value for the final estimate of the evolution variance W_t. This is what is referred to as discounted evolution variance.

2.5. **MCMC Diagnostics.** A critical issue for users of MCMC methods is how to determine when it is safe to stop sampling and use the samples to estimate characteristics of the distribution of interest. In this section, we detail the MCMC diagnostic tools used. The convergence diagnostics of [24] is used to compare values in the early part of the Markov chain to those in the latter part of the chain in order to detect failure of convergence. The statistic is constructed as follows: Two sub-sequences of the Markov chain θ are taken out, with $\theta_1^t : t = 1, ..., n_1$ and $\theta_2^t : t = n_a, ..., n$ where $1 \leq n_1 \leq n_a < n$.
Let $n_2 = n - n_a + 1$ and define $\bar{\theta}_1 = \frac{1}{n_1}\sum_{t=1}^{n_1} \theta^t$ and $\bar{\theta}_2 = \frac{1}{n_2}\sum_{t=na}^{n} \theta^t$ .Geweke test statistics was used to test whether the mean estimates have converged by comparing means from the early and latter part of the Markov chain. Assuming the ratios $\frac{n_1}{n}$ and $\frac{n_2}{n}$ are fixed,

$\frac{n_1+n_2}{n} < 1$, then the following statistic converges to standard normal distribution as n approaches ∞ we have

$$Z_n = \frac{\bar{\theta}_1 - \bar{\theta}_2}{\sqrt{\hat{s_1}(\theta)/n_1 + \hat{s}_2(\theta)/n_2}} \tag{33}$$

where $\hat{s_1}(\theta)$ and $\hat{s_2}(\theta)$ respresents spectral density estimates at zero frequencies. This is a two-sided test and large absolute value $Z-score$ indicates rejection of the null hypothesis of non-stationarity. Effective sample size on the other hand, relates to autocorrelation and measures mixing of the Markov chain. Most often, large discrepancy between the effective sample size and the simulation sample size indicates poor mixing. Effective Sample Size (ESS) is defined as

$$ESS = \frac{n}{\eta} = \frac{n}{1 + 2\sum_{k=1}^{\infty} \rho_k(\theta)} \tag{34}$$

where n is the total sample size and $\rho_k(\theta)$ is the autocorrelation at lag k for θ. The quantity η is autocorrelation time. The Bayesian process for estimating it is to first find a cut off point k after which the autocorrelations are very close to zero and then sum all the ρ_k to that point. The cut off point k is such that $\rho_k < 0.01$ or $\rho_k < 2s_k$ where s_k is the standard deviation defined as

$$s_k = 2\sqrt{(}\frac{1}{n}(1 + 2\sum_{j=1}^{k-1} \rho_k^2(\theta))) \tag{35}$$

Very low values of ESS often indicate poor mixing of the Markov chain while computing the parameters.

3. Empirical Applications and Data Analyses

The application of the methods in this study is first undertaken with a simulation exercise of the dynamic Bayesian regression model to examine its estimation against the possibility of determining structural changes and corresponding best choice of models and discount values.

3.1. **Simulation Study.** In our simulation study, a simulated response variable (y_{ti}) was regressed on nine simulated covariates ($x_{ti}, i = 1, ..., 9$) with various adjustments to the simulated intercepts, slopes and variances of the model against the possibility of various structural changes, with emphasis on the role of discount values in model selection and estimation.

3.1.1. *Data Generating Process.* Using the R statistical software package, we simulated the response variable y_t as a simple random walk, or Brownian motion:

$$y_t|y_{t-1} \sim N(y_{t-1}, \sigma_y^2). \tag{36}$$

where the initial value y_1 is chosen as 2. The simulation experiment involve time varying intercepts, slopes and variances.

Let $A = diag(a_t)$ and $\underline{b} = \{b_t\}$, we allowed $\underline{\epsilon}$ to have a time-varying variance, giving

$$x_j = A\underline{y} + \underline{b} + \underline{\epsilon} \tag{37}$$

for $j = 1, ..., 9$ and $\underline{\epsilon} \sim MVN(\underline{0}, \Sigma_j)$ where Σ_j is a diagonal covariance matrix with arbitrary variances for each point in time.

$\sigma_y^2 = 2$ having time-varying slopes for sample sizes varying from 20, 30, 40, 50, 60, 70, 80, 90 and 100. The models were run with 12000 iterations each with a burn-in of 2000 samples. In other words, we drew M = 10, 000 samples after the initial 2,000 samples were discarded. Figure 1 shows the plot of the simulated data considered for the time varying intercepts, slopes and variances. In the simulation experiment, we adjust covariates 2, 5 and 9 to follow increasing order of sudden jumps in order to detect their peculiar behaviour with respect to the choice of the best models and discount values. It was observed that all the models converged well as the absolute value of their Geweke Convegence Diagnostics (GCD) fell below the 1.96 threshold with sufficient Effective Sample Sizes (ESS) generated by our algorithm. Figure 2 is a plot of estimated slopes obtained from simulated data. It shows that the method used here is able to detect sudden jumps, regime shifts and structural breaks as characterized in the trend trajectory shown in the diagram. Figure 3 shows clearly that the chain mixes well.

3.2. **More results and insights from the simulation experiments.** We consider moving dynamic regressions in the spirit of [25] to test the constancy of the behaviour of the estimated parameters along the years for time-varying intercepts, slopes and variances in the simulation experiment. Tables 1 to 9 represent simulation experiments with varying sample sizes such as n= 20, 30, . . . , 100 in order to verify if optimal discount values respond to variations in the input coefficients of the models. In Table 1 (n=20) x_9 is the best regressor (with the minimum Mean Squared Predicted Error (MSPE) value of 0.116), while x_6 is the regressor with the

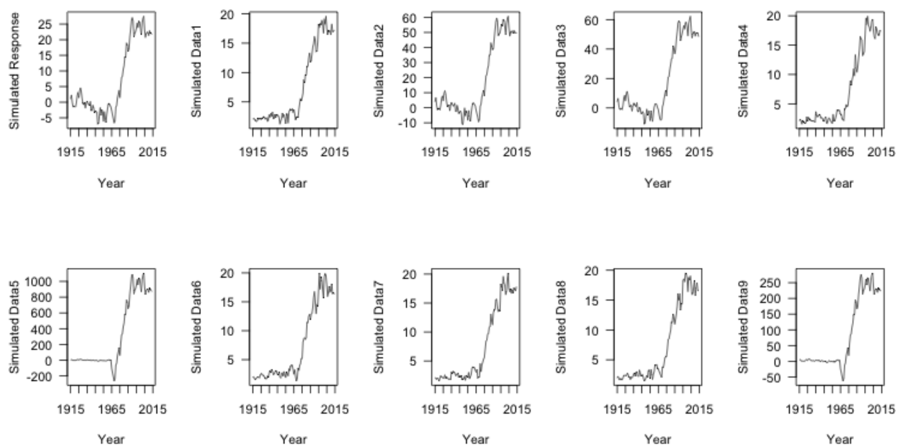

FIGURE 1. Plot of Simulated Data

highest volatility rate (with the minimum λ of value 0.340). In Table 2 (n=30) x_9 is the best regressor (with the minimum MSPE of value 0.180), x_6 is also the regressor with the highest volatility rate (with the minimum λ of value 0.520). In Table 3 (n=40) x_9 is the best regressor (with the minimum MSPE of value 0.180), x_6 is the regressor with the highest volatility rate (with the minimum λ of value 0.540).

Table 1. Dynamic Regression, n=20

Model	Regressor	MSPE	V	GCD	ESS	λ
1	x_1	5.042	1.603	0.993	3122	0.460
2	x_2	0.229	0.031	-1.380	9042	0.970
3	x_3	5.175	1.921	-1.142	4981	0.610
4	x_4	3.972	1.382	-2.732	3413	0.530
5	x_5	0.309	0.058	0.684	8777	0.980
6	x_6	2.921	122.565	0.996	6685	0.340
7	x_7	1.457	0.586	1.031	8369	0.720
8	x_8	6.377	2.560	2.003	3649	0.480
9	x_9	**0.116**	83.071	1.016	10139	0.570

The plots of the estimated time-varying slopes of simulated data (not shown) indicates that lower discount values are associated with more volatile relationship between the response and predictor (regressor) variables.

Table 2. Dynamic Regression, n=30

Model	Regressor	MSPE	V	GCD	ESS	λ
1	x_1	6.432	2.789	-1.128	5599	0.660
2	x_2	0.232	0.053	0.939	4007	0.580
3	x_3	6.412	2.135	-1.856	5430	0.660
4	x_4	5.510	1.456	-1.043	4570	0.640
5	x_5	0.225	0.056	1.023	5156	0.650
6	x_6	3.216	0.912	0.670	3299	0.520
7	x_7	5.133	1.224	-0.225	4071	0.560
8	x_8	6.640	2.891	0.892	5054	0.610
9	x_9	**0.180**	0.033	0.785	4980	0.650

Table 3. Dynamic Regression, n=40

Model	Regressor	MSPE	V	GCD	ESS	λ
1	x_1	5.433	2.396	0.705	5279	0.660
2	x_2	0.192	0.047	-0.998	4750	0.630
3	x_3	5.244	1.791	-0.505	5751	0.680
4	x_4	4.550	1.222	-0.565	4762	0.630
5	x_5	0.181	0.050	0.439	5794	0.690
6	x_6	2.935	0.925	-0.701	3684	0.540
7	x_7	5.072	1.608	0.741	4415	0.600
8	x_8	5.691	2.590	-0.792	5212	0.620
9	x_9	**0.180**	0.045	-1.023	4450	0.590

In Table 4 (n=50), x_9 is the best regressor (with the minimum MSPE of value 0.229), x_5 is the regressor displaying the highest volatility rate (with the minimum of value 0.560). In Table 5 (n=60), x_6 is the best regressor (with the minimum MSPE of 3.12), x_9 is the regressor with the highest volatility rate (with the minimum λ of value 0.41). In Table 6 (n=70), x_9 is the best regressor (with the minimum MSPE of 2.90), and the highest volatility rate (with the minimum λ of value 0.42).

Table 4. Dynamic Regression, n=50

Model	Regressor	MSPE	V	GCD	ESS	λ
1	x_1	4.626	781.082	0.983	8573	0.670
2	x_2	0.312	2.842	0.997	6653	0.640
3	x_3	4.630	84.824	1.009	7573	0.670
4	x_4	4.067	1.151	0.732	5035	0.640
5	x_5	0.275	88.484	1.015	7468	0.560
6	x_6	3.170	65.715	1.050	7094	0.570
7	x_7	4.389	1.480	1.028	6623	0.620
8	x_8	4.944	2.335	-0.434	5219	0.640
9	x_9	**0.229**	0.054	-0.735	5105	0.620

In Table 7 (n=80), x_9 is the best regressor (with the minimum MSPE value of 2.53), and also the highest volatility rate (with the minimum λ of value 0.42) . In Table 8 (n=90), x_9 is the best regressor (with the minimum MSPE value of 2.25), and the highest volatility rate (with the minimum λ of value 0.41). In Table 9 (n=100), x_2 is the best regressor (with the minimum MSPE value of 0.09), x_9 is the covariate with the lowest Discount Factor indicating highest volatility rate (with the minimum λ value of 0.26). Form all the tables, it was seen that in all but one of the models, the model involving x_9 was the best with the minimum mean square prediction error. It also has the highest frequency of being the regressor with lowest discount factor (in 5 tables). It is so because x_9 has the highest rate of sudden jump in the simulation excercise. Some new insights can be gathered here. We observe that the use of discount values is both parsimonious and efficient and plays an important role in fast convergence of the MCMC technique used. Also, from the experiment, we notice that dynamic linear models respond well to sudden jumps, regime shifts and structural changes and generally, we observe that simulated data with the same seasonal behaviour have the same range of discount values. Generally, there is also decreased flunctuations in discount values as the sample size increases. Notice that similar simulated covariates exhibit the same range of discount values as revealed in Tables 1-9.

Table 5. Dynamic Regression, n=60

Model	Regressors	MSPE	V	GCD	ESS	λ
1	x_1	3.95	8017.79	1.01	10477	0.67
2	x_2	3.54	0.23	-0.05	2540	0.43
3	x_3	4.35	1.53	0.67	5906	0.68
4	x_4	3.69	1.19	-0.02	5086	0.64
5	x_5	3.59	0.25	1.31	2813	0.44
6	x_6	3.12	1.16	1.67	4693	0.59
7	x_7	4.43	1.62	0.94	5699	0.67
8	x_8	4.35	2.08	-0.28	5389	0.65
9	x_9	**3.39**	0.20	0.62	2577	0.41

Table 6. Dynamic Regression, n=70

Model	Regressors	MSPE	V	GCD	ESS	λ
1	x_1	4.13	1.73	-1.08	5371	0.66
2	x_2	3.03	0.19	0.99	2691	0.43
3	x_3	5.50	1.74	-0.04	5833	0.67
4	x_4	4.00	9.53	1.01	9868	0.62
5	x_5	3.07	0.21	0.95	3651	0.44
6	x_6	3.64	1.22	-0.63	4514	0.58
7	x_7	4.60	233.20	1.01	8468	0.65
8	x_8	4.57	545.90	1.01	7097	0.63
9	x_9	**2.90**	0.19	1.04	2850	0.42

Table 7. Dynamic Regression, n=80

Model	Regressors	MSPE	V	GCD	ESS	λ
1	x_1	3.91	1.68	1.11	6486	0.67
2	x_2	2.64	1.80	1.00	3559	0.44
3	x_3	5.01	7.34	1.03	7290	0.67
4	x_4	3.79	1.41	1.02	5861	0.63
5	x_5	2.69	0.18	1.55	2872	0.44
6	x_6	3.40	1.16	0.67	4374	0.59
7	x_7	4.39	1.50	1.62	5203	0.65
8	x_8	4.34	2.03	0.99	6476	0.65
9	x_9	**2.53**	0.15	0.57	2826	0.42

Table 8. Dynamic Regression, n=90

Model	Regressors	MSPE	V	GCD	E SS	λ
1	x_1	3.56	1.51	-0.60	5840	0.67
2	x_2	2.35	30.91	1.00	3401	0.44
3	x_3	4.72	2.26	1.03	7281	0.67
4	x_4	3.61	1.31	-0.16	5366	0.64
5	x_5	2.39	4.85	1.01	3618	0.44
6	x_6	3.26	2.98	1.02	5156	0.59
7	x_7	4.31	1.46	1.43	5455	0.65
8	x_8	4.06	2.04	0.99	6256	0.66
9	x_9	**2.25**	0.13	-0.62	2517	0.41

Table 9. Dynamic Regression, n=100

Model	Regressor	MSPE	V	GCD	ESS	λ
1	x_1	2.52	0.93	-0.12	4541	0.58
2	x_2	0.09	7.60	1.01	9023	0.73
3	x_3	0.10	0.05	0.27	6301	0.70
4	x_4	2.62	2.15	0.98	6894	0.58
5	x_5	24.23	0.07	-0.13	8470	0.81
6	x_6	2.45	1.00	-0.55	5369	0.63
7	x_7	3.60	1.65	0.99	9109	0.78
8	x_8	2.25	0.61	0.38	4900	0.60
9	x_9	**1.51**	0.02	1.08	1281	0.26

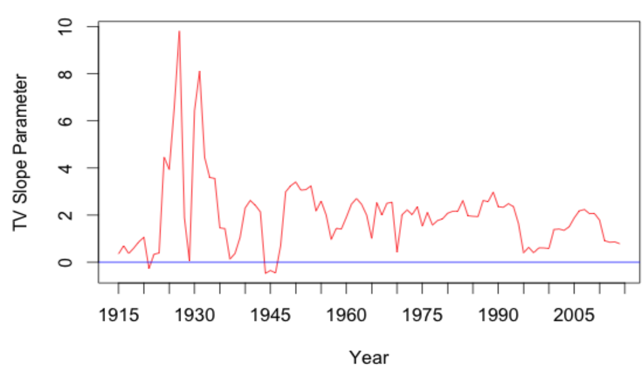

FIGURE 2. Time-Varying Slope Parameters: Simulated Data

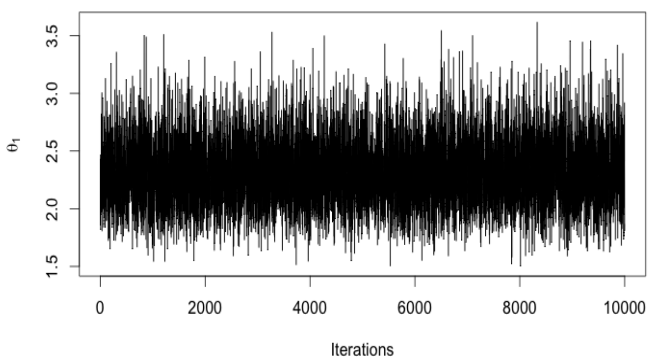

FIGURE 3. Trace Plot: Simulated Data.

3.3. **An Econometric Application.** This section involves an econometric application involving the predictors of Nigerian Crude Oil export using the proposed model and method outlined in Section 3. The data used in this research are Nigerian economic indicators sourced from the Central Bank of Nigeria (www.cbn.org). The data includes annual time-series data on Nigerian Oil-Export , GDP, Capital Expenditure, Consumer Price Index, Exchange Rate and Lending Rate of the pre-global recession period (1960-2009). GDP and export data were logged before analysis.
First, we estimate the Bayesian dynamic model to examine the predictive effect of each macroeconomic variable on oil-export in the presence of discounted evolution variance. It was discovered that lending rate performed better than other variables in the study in terms of predictive performance, given the low values of MSPE. However, we find that GDP performed better in predicting oil-export when combined with lending rate. In the tables showing the various model results (Tables 10 and 11), the discount values λ represents the level of fluctuation of the relationships between export and its predictors. More so, discount values tend to increase with the inclusion of additional regressors into the model as revealed in Table 11. High values of λ represents more smoothly changing fluctuation (volatility), while lower discount values represent more volatile relationships. Also, lower values of MSPE indicates a better predictive performance, while higher values indicate a lower predictive performance of the variables in the models. Table 10 shows

the dynamic regression of Oil-Export on selected related macroeconomic variables.

Table 10. Dynamic Linear Regression of Oil-Export on Selected Macroeconomic Variables

Model	Regressor	MSPE	V	GCD	ESS	λ
1	GDP	0.139	7.375	0.985	8422	0.56
2	CE	0.224	0.050	1.174	6981	0.69
3	CPI	0.155	0.029	0.360	5131	0.59
4	EXRT	0.734	0.039	-0.506	5316	0.61
5	LR	0.129	0.027	1.433	2284	0.32

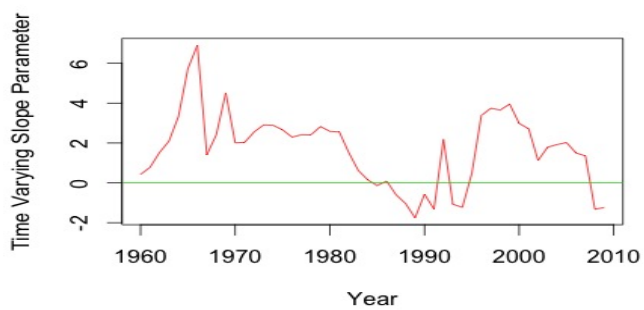

FIGURE 4. Time-Varying Slopes of Oil Export vs GDP, $\lambda = 0.56$.

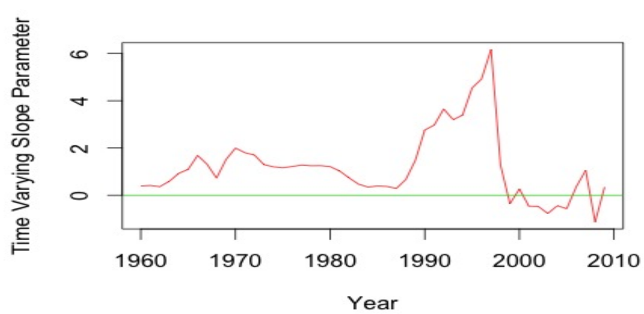

FIGURE 5. Time-Varying Slope of Oil Export vs Capital Expenditure, $\lambda = 0.69$

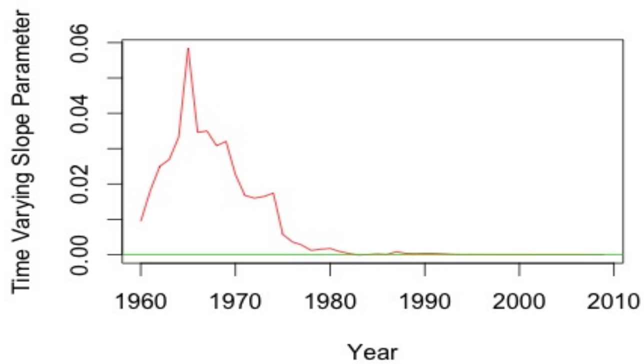

FIGURE 6. Time-Varying Slopes of Oil Export vs Consumer Price Index, $\lambda = 0.59$

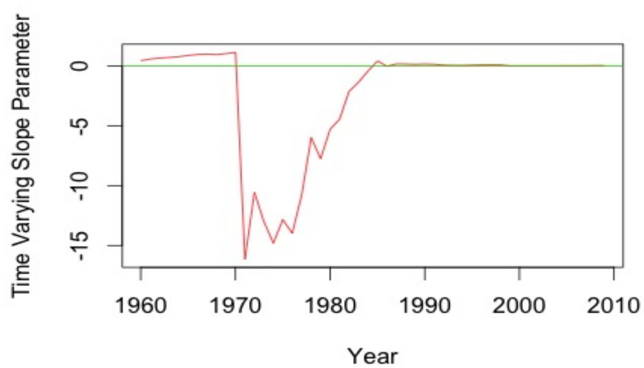

FIGURE 7. Time-Varying Slopes of Oil Export vs Exchange Rate, $\lambda = 0.61$

Table 11. Dynamic Regression of Oil Export on Various Economic Variables

Model	Regressor	MSPE	V	GCD	ESS	λ
1	LR +GDP	0.122	0.014	1.164	2007	0.45
2	LR +CE	0.313	0.089	-1.012	7141	0.82
3	LR +CPI	0.198	0.027	1.039	2996	0.54
4	LR +EXRT	1.488	0.619	0.574	10754	0.99

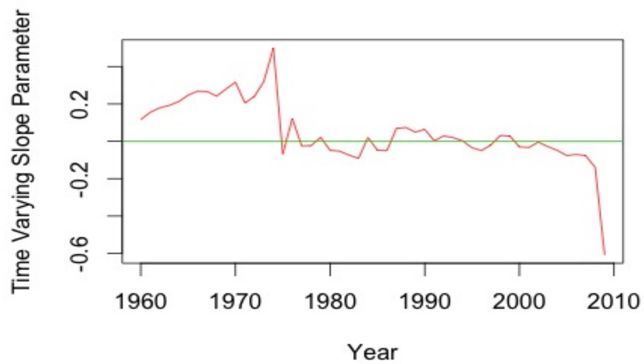

FIGURE 8. Time-Varying Slopes of Oil Export vs Lending Rate, $\lambda = 0.32$

4. Discussion of Results and Policy Recommendations

There are quite a number of interesting results from this application. Table 10 shows the dynamic regression of oil export on its predictors. It was discovered that lending rate, among others, has the lowest predictive error of 0.129. This implies that increased oil export led to greater revenue for the Nigerian government compared to the pre-oil era. The aftermath of this was an expansion of money supply. Having more liquidity in the economy has a depressing effect on the lending rate. The lending rate is the cost of capital. As it reduces, it encourages more investment activities either for domestic or export. It is therefore very pertinent for the monetary authority in Nigeria to ensure the stability of the lending rate in order to stimulate investment and agricultural(non-oil) export business in Nigeria in the post-pandemic era.

In a further analysis, we find that different combinations of lending rate with GDP, capital expenditure, CPI and the exchange rate gave better insight and line of policy decisions. From Table 11, lending rate combined with GDP gave a minimum MSPE of 0.122 with respect to oil export. The implication of this is that lower lending rates accessible to the investors will stimulate the economy. The lower the value of MSPE, the more the contribution of the variable in the model and the higher its predictive power. Growth in GDP implies increased national income which has a stimulating effect on oil export, therefore, necessitating the inclusion of GDP in the model.

From the foregoing, it is very important that the government of Nigeria through the monetary authority formulate polices that will help eliminate erratic fluctuations in the lending rate. In addition, the time-varying slope parameter depicts some levels of associations among the variables over the years. For instance, the time-varying slope of oil export and GDP in Figure 4 shows positive association over the years 1960 to early 1980. The association was strongest in the 1960s as can be seen. It however declined sharply towards the late 1960s, thereafter fluctuating and eventually declining sharply. This can be attributed to decreased demand for oil-export from Nigeria in the world market. The picture changed completely around the mid-1980s and 1990s at which oil export correlates negatively with GDP. The relationship turned positive between the mid 1990 and late 2009 due to an improvement in the demand for oil export from Nigeria during the period. Trends depicted in Figure 5 shows that oil export and capital expenditure remained positively correlated over the stretch of 1960 to late 1990s. For instance, capital expenditure in the form of oil depots and other viable resources will help to boost oil export. It was around mid 1990s during the dictatorial military regime in Nigeria that this relationship was strongest. The government should implement policies that would invest in more capital expenditures in the post-COVID-19 era.

It can be seen from Figure 6 that the consumer price index which is a measure of the relative change in the price level of goods and services positively correlates with oil export between 1960 to early 1980 after which it declined steadily. Increased oil export generates more income for the government which through its multiplier effect will stimulate aggregate demand in the economy with its consequent effect on the CPI. The pattern of exchange rate with respect to oil export has been rather volatile with only positive relationship in the 1960s after which there was a sharp decline around early 1970s as contained in Figure 7. The negative relationship continued over a long period till early 1980s to 2009 at which it turned zero. The government through the central bank should have helped to monitor the exchange rate appropriately in order to maintain its stability given its influence on the proceeds from oil export. The relationship between oil export and lending rate was not stable over the study period. The result is displayed in Figure 8 where the time- varying parameter estimates appears to be erratic. Given the results of its MSPE, the lending rate appears crucial. The monetary authority of

Nigeria should formulate policies towards stabilizing inflation and the govenment should focus more on non-oil (agricultural) export after the pandermic as suggested by [17]. Discounting played crucial role in reducing the times (not shown) and intensity of all these models.

5. Concluding Notes

This paper has examined the problem of estimating dynamic state space models in the presence of discounted evolution variance using simulated and economic data with several newly revealed insights. We have presented a class of modular dynamic time series regression models that can accommodate several different dynamics with the ability to estimate time-varying parameters. The proposed Bayesian analysis using MCMC method with embedded RFFBS allows for a full account of uncertainty in the model and can be performed with moderate computational resources due to the fast convergence of the chain as a result of discounting of the state variance. It was revealed that the estimated model is able to detect regime shifts, structural changes and sudden jumps in the historical predictors of crude oil export studied while portending some policy implications for the post-pandemic era. The procedures developed and presented in this paper will be useful in many areas of research where estimation of time-varying parameters in dynamic models are of interest. There are several promising directions for further research and policy applications of the class of models presented in this article. One of such is the generalization to simultaneous equations model to enable multivariate response observations in space and time. An example of such is the Bayesian Functional ANOVA models where the response is naturally a function of space and time. Another possibility would be to test and compare the forecasting performance and goodness of fit of our model in relation with other models including linear ordinary least squares model and a machine learning model like the artificial neural network model which has recently gained tremendous popularity for forecasting non-linear time series data. Works in these areas are in progress with a plethora of possible future policy implications.

ACKNOWLEDGEMENTS

Many thanks to two anonymous erudite reviewers whose prompt and constructive comments led to the improvement of the first version of this article. All glory be to God Almighty!

References

[1] G. Petris, S. Petrone, and P. Campagnoli, *Dynamic linear models with R*. Springer, 2009.

[2] J. Fúquene, M. Álvarez, and L. R. Pericchi, "A robust bayesian dynamic linear model for latin-american economic time series:"the mexico and puerto rico cases"," *Latin American Economic Review*, vol. 24, no. 1, pp. 1–17, 2015.

[3] R. Ravines, A. M Schmidt, and H. S Migon, "Revisiting distributed lag models through a bayesian perspective," *Applied Stochastic Models in Business and Industry*, vol. 22, no. 2, pp. 193–210, 2006.

[4] G. U. Yule, "Why do we sometimes get nonsense-correlations between time-series?–a study in sampling and the nature of time-series," *Journal of the royal statistical society*, pp. 1–63, 1926.

[5] J. Ojo, "On time series models and prediction of deposits and loans of rural branches of commercial banks in nigeria," *Journal of Science and Multidisciplinary Research*, vol. 5, no. 2, 2013.

[6] S. Almon, "The distributed lag between capital appropriations and expenditures," *Econometrica: Journal of the Econometric Society*, pp. 178–196, 1965.

[7] H. Wold, "A study in the analysis of stationary time series," 1938.

[8] P. J. Harrison and C. F. Stevens, "Bayesian forecasting," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 205–247, 1976.

[9] M. West and P. Harrison, *Bayesian Forecasting and Dynamic Models*, 2nd ed. New York: Springer-Verlag, 1997.

[10] J. Nakajima, M. Kasuya, and T. Watanabe, "Bayesian analysis of time-varying parameter vector autoregressive model for the japanese economy and monetary policy," *Journal of the Japanese and International Economies*, vol. 25, no. 3, pp. 225–245, 2011.

[11] T. Doh and M. Connolly, *The state space representation and estimation of a time-varying parameter VAR with stochastic volatility*. Springer, 2013.

[12] K. Cui and D. B. Dunson, "Generalized dynamic factor models for mixed-measurement time series," *Journal of Computational and Graphical Statistics*, vol. 23, no. 1, pp. 169–191, 2014.

[13] V. Soloviev, V. Saptsin, and D. Chabanenko, "Markov chains application to the financial-economic time series prediction," *arXiv preprint arXiv:1111.5254*, 2011.

[14] G. E. Primiceri, "Time varying structural vector autoregressions and monetary policy," *The Review of Economic Studies*, vol. 72, no. 3, pp. 821–852, 2005.

[15] A. H. Sarris, "A bayesian approach to estimation of time-varying regression coefficients," in *Annals of Economic and Social Measurement, Volume 2, number 4*. NBER, 1973, pp. 497–520.

[16] L. A. Gil-Alana, R. Gupta, O. E. Olubusoye, and O. S. Yaya, "Time series analysis of persistence in crude oil price volatility across bull and bear regimes," *Energy*, vol. 109, pp. 29–37, 2016.

[17] O. O. Awe, D. M. Akinlana, O. S. Yaya, and A. O., "Time series analysis of the behaviour of import and export of agricultural and non-agricultural

goods in west africa: A case study of nigeria." *Agris On-line Journal of Economics and Informatics*, vol. 10, no. 2, pp. 15–22, 2018.
[18] B. Samson and O. Mercy, "On the time series modelling of crude oil exportation in nigeria," *European Journal of Academic Essays*, vol. 2, no. 3, pp. 1–11, 2015.
[19] K. Ayinde and H. Abdulwahab, "Modeling and forecasting nigerian crude oil exportation: Seasonal autoregressive integrated moving average approach," *International Journal of Science and Research*, vol. 2, no. 12, pp. 245–249, 2013.
[20] J. M. Marin and C. Roberts, *Bayesian Essentials with R.* Springer-Verlag New York (Springer Texts in Statistics. 2nd Edition), 2014.
[21] O. O. Awe, I. Crandell, and A. A. Adepoju, "A time varying parameter state-space model for analyzing money supply-economic growth nexus," *Journal of Statistical and Econometric Methods*, vol. 4, no. 1, pp. 73–95, 2015.
[22] A. Gelman, "Parameterization and bayesian modeling," *Journal of the American Statistical Association*, vol. 99, no. 466, 2004.
[23] O. O. Awe and A. A. Adepoju, "Modified recursive bayesian algorithm for estimating time-varying parameters in dynamic linear models," *Statistics in Transition-New Series*, vol. 19, no. 2, pp. 239–258, 2018.
[24] J. Geweke, "Bayesian treatment of the independent student-t linear model," *Journal of Applied Econometrics*, vol. 8, no. S1, pp. S19–S40, 1993.
[25] R. L. Brown, J. Durbin, and J. M. Evans, "Techniques for testing the constancy of regression relationships over time," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 149–192, 1975.

DEPARTMENT OF MATHEMATICAL SCIENCES, ANCHOR UNIVERSITY LAGOS, NIGERIA;
DEPARTMENT OF STATISTICS, FEDERAL UNIVERSITY OF BAHIA, SALVADOR, BAHIA, BRAZIL.
E-mail addresses: olawaleawe@gmail.com, oawe@aul.edu.ng

DEPARTMENT OF STATISTICS, UNIVERSITY OF IBADAN, IBADAN,NIGERIA
E-mail address: pojuday@yahoo.com