DYNAMIC LINEAR REGRESSION BY BAYESIAN AND BOOTSTRAPPING TECHNIQUES

A. ADEDAYO ADEPOJU
Department of Statistics, UNIVERSITY OF IBADAN, NIGERIA
E-mail: pojuday@yahoo.com

TAYO. P. OGUNDUNMADE
Department of Statistics, UNIVERSITY OF IBADAN, NIGERIA
E-mail: ogundunmadetayo@yahoo.com

ABSTRACT

Estimation of a dynamic linear regression is said to be of importance as events are measured on past events. However, estimation of dynamic models using the OLS is inefficient since it cannot produce the least variance and disregarding this problem would potentially lead to severe statistical problems. Therefore, the main objective of the study is to employ the bootstrap technique and Bayesian method in estimating the parameters of a dynamic regression model and compare their performances using standard deviation, absolute bias and mean square error of the estimates. The residual resampling technique is used for the bootstrap approach and the Normal-gamma model for the Bayesian approach. The results showed that the bootstrap technique outperformed the Bayesian method with lower standard error values. The bootstrap method also displayed an asymptotic property.

Key words: Dynamic Model, Bootstrap Method, Bayesian Estimator, Absolute Bias, Mean Square error.

Regresión lineal dinámica por técnicas bayesianas y bootstrapping

RESUMEN

Se dice que la estimación de una regresión lineal dinámica es importante a medida que los eventos se miden en eventos pasados. Sin embargo, la estimación de los modelos dinámicos que utilizan el MCO es ineficiente, ya que no puede producir la menor variación y el hecho de no tener en cuenta este problema podría conducir a problemas estadísticos graves. Por lo tanto, el objetivo principal del estudio es emplear la técnica bootstrap y el método bayesiano para estimar los parámetros de un modelo de regresión dinámica y comparar sus rendimientos utilizando desviación estándar, sesgo absoluto y error cuadrático medio de las estimaciones. La técnica de remuestreo residual se utiliza para el enfoque bootstrap y el modelo de gamma normal para el enfoque bayesiano. Los resultados mostraron que la técnica de arranque superó al método Bayesiano con valores de error estándar más bajos. El método de arranque también mostró una propiedad asintótica.

Palabras clave: modelo dinámico, método Bootstrap, estimador bayesiano, sesgo absoluto, error cuadrático medio.

Clasificación JEL: C5

1.   INTRODUCTION

A linear regression is used mainly to model a relationship between two or more than two variables. In general, the most popular method used for regression is Ordinary Least Squares (OLS) for its ease and simplicity. When considering a dynamic model, the lag of the dependent variable is used as independent variable or the lag of one or more of the independent variables is also considered as another independent variable in the same model. In regression scenario, most commonly used bootstrap methods are residual bootstrap and pair bootstrap. Several researchers have introduced various remedial (also known as classical bootstrap methods).

Efron (2012), used the parametric bootstrap to carry out Bayesian inference calculations. Two main points were made: that in the comparatively limited set of cases where bootstrap methods apply, they offer an efficient and computationally straightforward way to compute posterior distributions and estimates, enjoying some advantages over Markov chain techniques; and, more importantly, that the parametric bootstrap helps connect Bayes and frequentist points of view.

Thai et.al, (2013), used simulation to evaluate bootstrap methods proposed for linear mixed-effect models. They also compared the results obtained by maximum likelihood (ML) and restricted maximum likelihood (REML). Their simulation studies evidenced the good performance of the case bootstrap as well as the bootstraps of both random effects and residuals. However, the bootstrap methods which resample only the residuals and the bootstraps combining case and residuals performed poorly.

Zahari et.al (2014) employed combination of a bootstrapping technique and robust methods in ridge regression model for data with simultaneous problems of multicollinearity and multiple outliers. They employed the fixed - x resampling technique for robust ridge regression. The proposed method was expected to reduce the effect of the problems to the estimation results by producing lower standard error values as compared to the existing methods. The results showed that the proposed technique produced better parameter estimates with lower standard error values.

Zhang and Worthington (2015) demonstrated methods for constructing confidence interval for scope economies to allow researchers to draw inferences from estimated economies of scope. They reviewed the common approaches such as delta method or bootstrap adopted by previous studies. In contrast of the above approximation methods, they also proposed an alternative method, Bayesian approach, to produce full predictive distribution for this measure with posterior distribution. To demonstrate these three approaches, they used a balanced panel data including 37 Australian public universities over the period 2003-12. All three approaches used a quadratic cost function with two outputs. Estimates of scope economies were calculated with the sample data and estimated parameters from the model. Results showed that the Bayesian approach gives the most precise (the least standard deviations) among all approaches.

Other researches on the bootstrap methods and dynamic models are; Hansen (1999), Davison and Hinkley (1997), Efron and Tibshirani (1993), Hall (1986, 1989, 1992), Li and Maddala (1996}, Horowitz (1997) and Berkowitz and Kilian (2000).

The remainder of the paper is organised as follows: section 2 discusses the methodology followed by the data generation process in section 3. Analysis and the presentation of results are presented in section 4 while the concluding remarks are given in section 5.

2.   MATERIALS AND METHODS

Consider a dynamic model given as

$$y_t = \delta y_{t-1} + \beta_1 + \beta_2 x_t + e_t \tag{1}$$

Where,

$y_t =$ is the dependent variable for the tth period.

$y_{t-1} =$ is the lagged of the dependent variable.

$\delta =$ is the coefficient of the lagged dependent variable.

$x_t =$ is the exogenous variables.

$e_t =$ is the error term.

2.1. Bayesian Approach

In statistics, Bayesian linear regression is an approach in which the statistical analysis is undertaken within the context of Bayesian inference. When the regression model has errors that have a normal distribution, and if a particular form of prior distribution is assumed, explicit results are available for the posterior distributions of the model's parameters. Consider a dynamic linear model in equation (1) above, normal distribution is used for its likelihood and gamma distribution for the conjugate prior in this study.

Equation (1) above can be written more compactly as:

$$y_t = \theta_i X_t + e_t, \qquad e_t \sim N(0, h^{-1} I) \tag{2}$$

Where, $X_t = (y_{t-1}, 1, x_t)$, $\theta_i = (\delta, \beta_1, \beta_2)^I$, I is identity matrix.

Hence, the transformed model given in (2) above is identical to the normal linear regression model. The vector of disturbance terms e_t, is assumed to be a Normal distribution with mean 0 and covariance σ^2, $N(0, h^{-1})$ where $h = \sigma^{-2}$, the error precision.

2.1.1. The Likelihood Function

Using the definition of the multivariate Normal density, the likelihood function for the model can be written as:

$$P(y \mid \beta, h) = \frac{h^{N/2}}{(2\pi)^{N/2}} \left\{ exp\left[-\frac{h}{2} (y - X\hat{\theta})'(y - X\hat{\theta}) \right] \right\} \left\{ h^{v/2} exp\left[-\frac{hv}{2s^{-2}} \right] \right\} \tag{3}$$

where,

$$v = N - k$$
$$\hat{\theta} = (X^I X)^{-I} X^I Y$$

is the ordinary least squares estimator and

$$s^2 = \frac{(y - X\hat{\theta})'(y - X\hat{\theta})}{v}$$

is the variance for the error (Mean Square Error).

2.1.2. The Priors for Normal Linear Regression Models

Since the likelihood function of models determines the structure or distribution of the prior especially for easier interpretations and computations therefore, the natural conjugate prior is Normal-Gamma density. Thus, if we elicit a prior for θ conditional on h of the form:

$$\theta | h \sim N(\underline{\theta}, \underline{V})$$

Given as

$$P(\theta | h) = \frac{\underline{V}^{\frac{1}{2}}}{(2\pi)^{N/2}} \left\{ exp\left[-\frac{1}{2} (\theta - \underline{\theta})' \underline{V}^{-1} (\theta - \underline{\theta}) \right] \right\} \tag{4}$$

and a prior for h of the form:

$$h \sim G(\underline{s}^{-2}, \underline{v})$$
$$P(h) = \left\{ h^{v/2} exp\left[-\frac{hv}{2s^{-2}} \right] \right\} \tag{5}$$

then, the joint prior for the two parameters is given as:

$$\theta h = (\theta | h \times h) \sim NG(\underline{\theta}, \underline{V}, \underline{s}^{-2}, \underline{v})$$

Expressed as
Posterior = likelihood x Prior

$$P(\theta|h \times h) = \frac{\underline{V}^{\frac{1}{2}}}{(2\pi)^{N/2}}\left\{exp\left[-\frac{1}{2}(\theta - \underline{\theta})'\underline{V}^{-1}(\theta - \underline{\theta})\right]\right\} \times \left\{h^{\underline{v}/2}exp\left[-\frac{h\underline{v}}{2\underline{s}^{-2}}\right]\right\} \tag{6}$$

2.1.3. The Posterior

The Posterior will be obtained by multiplying both the likelihood and the prior distribution, that is, multiplying (3) and (6), we have

$$P(\theta, h|y) = \frac{h^{\frac{N}{2}}}{(2\pi)^{N/2}}\left\{exp\left[-\frac{h}{2}(y - X\hat{\theta})'(y - X\hat{\theta})\right]\right\}\left\{h^{v/2}exp\left[-\frac{hv}{2s^{-2}}\right]\right\} x \tag{7}$$

$$\frac{\underline{V}^{\frac{1}{2}}}{(2\pi)^{N/2}}\left\{exp\left[-\frac{1}{2}(\theta - \underline{\theta})'\underline{V}^{-1}(\theta - \underline{\theta})\right]\right\} \times \left\{h^{\underline{v}/2}exp\left[-\frac{h\underline{v}}{2\underline{s}^{-2}}\right]\right\}$$

The joint posterior density above in (7) does not take the form of any well-known density that is, it is not analytical. However, the conditionals of the posterior cannot be solved analytically but only through a posterior simulation method. It can be verified that the posterior of θ, conditional on h is multivariate Normal;

$$\theta|y, h \sim NG(\theta^*, V^*) \tag{8}$$

Where:

$$V^* = (\underline{V}^{-1} + hX'X)^{-1} \tag{9}$$

$$\theta^* = V^*(\underline{V}^{-1}\theta + X'X\hat{\theta}) \tag{10}$$

The posterior of h conditional on θ in the model is Gamma $h|y, \theta \sim G(s^{*-2}, v^*)$
Where,

$$v^* = v + N$$

and s^{*-2} is defined implicitly through:

$$v^*s^{*2} = vs^2 + [(y - X\hat{\theta})(y - X\hat{\theta})]^{-1} \tag{11}$$

2.1.4. The Posterior Means and Variances

By integrating out from (7), it gives the marginal posterior distribution for θ which is a multivariate t distribution given as:

$$\theta|h \sim N(\theta^*, s_*^2, V^*, \ v^*) \tag{12}$$

The mean,

$$E(\theta|h) = \theta^* = V^*X^{*'}y \tag{13}$$

and the variance,

$$Var(\theta|h) = \frac{v^*s_*^2}{v^* - 2}V^* \tag{14}$$

while posterior for the precision given the data follows the gamma distribution:

$$h/\theta \sim G(s^{*-2}, v^*) \tag{15}$$

then the mean,

$$E(h/\theta) = s^{*-2} \tag{16}$$

and the variance,

$$var(h/\theta) = \frac{2s^{*-2}}{v^*} \tag{17}$$

2.2. Bootstrap Approach

Bootstrapping is the practice of estimating properties of an estimator by measuring those properties when sampling from an approximating distribution. In the case where a set of observations can be assumed to be from an independent and identically distributed population, this can be implemented by constructing a number of resamples with replacement, of the observed dataset and of equal size of the observed dataset. There are different types of bootstrapping schemes; case resampling, Bayesian bootstrap, smooth bootstrap, resampling residuals, wild bootstrap, block bootstrap for time series etc.

This study employed the resampling residual technique. It involves fitting the model and bootstrapping the residuals with OLS.

(i)   Classical bootstrap approach with OLS (bootstrapping the residual)

- fit the model and obtain the residuals as $\hat{e} = y_i - g_i(\hat{\beta})$

- place probability $1/n$ at each $\hat{e}_i$ and sample the residual independently with replacement from $\hat{e}_1, \hat{e}_2, \ldots, \hat{e}_n$.

- generate bootstrap residuals e_i^* for $i = 1,2,\ldots,n$

- we then have a bootstrap data set $y_i^* = g_i(\hat{\beta}) + e_i^*$ $for\ i = 1,2,\ldots,n$

- for each bootstrap data set, we obtain $\hat{\beta}^*$

- the procedure is repeated B times and the covariance matrix for $\hat{\beta}$ is estimated; $= \frac{1}{B-1}\sum_{j=1}^{B}(\hat{\beta}_j^* - \hat{\beta}_*^*)(\hat{\beta}_j^* - \hat{\beta}_*^*)^T$,

- where $\hat{\beta}_j^*$ is the bootstrap estimate from the jth bootstrap sample and $\hat{\beta}_*^* = \frac{1}{B}\sum_{j=1}^{B}\hat{\beta}_j^*$

3.   DATA GENERATION PROCESS

The main task is the generation of stochastic dependent Y_t which is used in estimating the parameters of the model and lag Y_{t-1}.is obtained recursively. To achieve this, the following have to be assumed

(i) Values of the predetermined variable X_t

(ii) Values of the parameter, $\delta = 0.01, \beta_1 = 0.1, \beta_2 = 0.2$

The simulation of the error term e_t is generated from a normal distribution with mean 0 and variance 0.25. We proceed as follows:

a.   The sample size N is specified as N = 20, 30, 50 and 100.

b.   For all cases X_t are generated from a uniform (0,1)

c.   For each bootstrap method, B = 500, 1000 and 5000 bootstrapped random samples were generated. Replications using 500, 1000 and 5000 are employed for the Bayesian approach.

For easy comparison and evaluation of the performance of the estimators, the following criteria are used:

(i)   Mean values of the parameter estimates.

(ii)   Absolute Bias

(iii)   Standard deviation

(iv)   Mean square error

4.   ANALYSIS AND PRESENTATION OF RESULTS

Tables 1 and 2 below show the bootstrapping and the Bayesian results respectively. For the bootstrap, the result is displayed for the dynamic model considered using mean estimate, standard deviation, absolute bias and the mean square respectively, 500, 1000 and 5000 times of booting are considered at different sample sizes of 20, 30, 50 and 100. The same approach is used in Table 2 for the Bayesian result except that replication is used instead of boot. Replications of 500, 1000 and 5000 are also considered for easy comparison.

As the sample size increases, the standard deviation of all the parameter estimates decreases consistently. The values of the mean square error also show asymptotic behaviour. In all cases, as the number of bootstrap increases from 500 to 1000 the standard deviation decrease but slight increase is observed from 1000 to 5000. Interestingly, the absolute bias decrease consistently as the number of bootstrap increases.

Unlike the bootstrap method, the Bayesian method does not show any consistent asymptotic pattern. The standard deviation, absolute bias and mean square error (MSE) only decrease as the sample size increases in some cases. In most cases, least values of the standard deviation, absolute bias and MSE are produced when replication is 500.

Table 1

Bootstrapping Result

Sample size	Coefficient	Boot = 500				Boot = 1000				Boot = 5000			
		Estimate	Standard deviation	Abs Bias	MSE	Estimate	Standard deviation	Abs Bias	MSE	Estimate	Standard deviation	Abs Bias	MSE
n = 20	$\delta = 0.01$	0.011785	0.176062	0.001785	0.031001	0.012435	0.172081	0.002435	0.029618	0.012815	0.173923	0.002815	0.030257
	$\beta_1 = 0.1$	0.101079	0.071189	0.001079	0.005069	0.092144	0.070928	0.007856	0.005093	0.099828	0.071502	0.000172	0.005113
	$\beta_2 = 0.2$	0.198911	0.080795	0.001089	0.006529	0.19948	0.0812	0.000521	0.006594	0.19865	0.079113	0.00135	0.006261
n = 30	$\delta = 0.01$	0.00767	0.130871	0.00233	0.017133	0.007845	0.130074	0.002155	0.016924	0.0088	0.134671	0.0012	0.018138
	$\beta_1 = 0.1$	0.101533	0.055604	0.001533	0.003094	0.102062	0.053629	0.002062	0.00288	0.099782	0.054443	0.000218	0.002964
	$\beta_2 = 0.2$	0.198926	0.061738	0.001074	0.003813	0.197856	0.059662	0.002144	0.003564	0.200676	0.058643	0.000676	0.003439
n = 50	$\delta = 0.01$	0.013591	0.121091	0.003591	0.014676	0.012458	0.120908	0.002458	0.014625	0.009752	0.122894	0.000248	0.015103
	$\beta_1 = 0.1$	0.098451	0.04525	0.001549	0.00205	0.098658	0.044854	0.001342	0.002014	0.099811	0.045835	0.000189	0.002101
	$\beta_2 = 0.2$	0.19957	0.04378	0.000431	0.001917	0.200813	0.043205	0.000813	0.001867	0.20047	0.043103	0.00047	0.001858
n = 100	$\delta = 0.01$	0.008843	0.096757	0.001157	0.009363	0.00929	0.098937	0.00071	0.009789	0.008684	0.097031	0.001316	0.009417
	$\beta_1 = 0.1$	0.100583	0.035235	0.000583	0.001242	0.100389	0.035636	0.000389	0.00127	0.100369	0.036	0.000369	0.001296
	$\beta_2 = 0.2$	0.197395	0.035545	0.002605	0.00127	0.198231	0.037056	0.001769	0.001376	0.199958	0.036795	4.23E-05	0.001354

Table 2
Bayesian Result

Sample size	Coefficient	Replication = 500				Replication = 1000				Replication = 5000			
		Estimate	Standard deviation	Abs Bias	MSE	Estimate	Standard deviation	Abs Bias	MSE	Estimate	Standard deviation	Abs Bias	MSE
n = 20	$\delta = 0.01$	0.051786	0.38406	0.041786	0.147502	0.031958	0.180622	0.021958	0.033106	0.036788	0.184567	0.026788	0.034782
	$\beta_1 = 0.1$	0.111601	0.161982	0.011601	0.026238	0.129316	0.412135	0.029316	0.170715	0.125129	0.40621	0.025129	0.165638
	$\beta_2 = 0.2$	0.211405	0.154235	0.011405	0.023919	0.233711	0.129667	0.033711	0.01795	0.232952	0.128845	0.032952	0.017687
n = 30	$\delta = 0.01$	0.013435	0.330519	0.003435	0.110989	0.010197	0.087936	0.000197	0.007733	0.011015	0.118211	0.001015	0.013975
	$\beta_1 = 0.1$	0.14445	0.130298	0.04445	0.017112	0.13411	0.143078	0.03411	0.021635	0.065163	0.150263	0.034837	0.023792
	$\beta_2 = 0.2$	0.1862	0.129563	0.0138	0.016917	0.163552	0.120687	0.036448	0.015894	0.21675	0.140603	0.01675	0.02005
n = 50	$\delta = 0.01$	0.036558	0.283756	0.026558	0.080529	0.039604	0.107907	0.029604	0.01252	0.006038	0.060045	0.003962	0.003621
	$\beta_1 = 0.1$	0.093183	0.107186	0.006817	0.013465	0.166817	0.281442	0.066817	0.083674	0.068634	0.27667	0.031366	0.07753
	$\beta_2 = 0.2$	0.202421	0.109592	0.002421	0.012201	0.235748	0.096159	0.035748	0.010524	0.275507	0.101578	0.075507	0.016019
n = 100	$\delta = 0.01$	0.024477	0.073017	0.014477	0.006037	0.021172	0.076708	0.011172	0.006009	0.014114	0.077035	0.004114	0.005951
	$\beta_1 = 0.1$	0.101246	0.022507	0.001246	0.050703	0.105791	0.221305	0.005791	0.049009	0.123053	0.222895	0.023053	0.050214
	$\beta_2 = 0.2$	0.208548	0.081244	0.008548	0.006606	0.286216	0.079045	0.086216	0.013681	0.289184	0.080959	0.089184	0.014508

4.1. Comparing Bootstrap and Bayesian approaches using Absolute Bias

Table 3 shows the comparison between the absolute biases produced using the two estimators under consideration. This explains how far the estimates are from the true values, the estimator with smaller absolute bias performs better than the other. The absolute biases produced using bootstrapping technique are consistently lower than those produced by the Bayesian at all levels of sample sizes. It implies that, using bootstrapping technique when considering a dynamic model will produce less absolute bias than using Bayesian technique. At boot = 5000 and for n>20, the bootstrap method produced the minimum absolute biases.

Table 3
Performance of Bootstrap and Bayesian approaches using Absolute Bias

Sample size	Coefficient	Bootstrapping			Bayesian		
		Boot			Replication		
		500	1000	5000	500	1000	5000
n = 20	$\delta = 0.01$	0.001785	0.002435	0.002815	0.041786	0.021958	0.026788
	$\beta_1 = 0.1$	0.001079	0.007856	0.000172	0.011601	0.029316	0.025129
	$\beta_2 = 0.2$	0.001089	0.000521	0.00135	0.011405	0.033711	0.032952
n = 30	$\delta = 0.01$	0.00233	0.002155	0.0012	0.003435	0.000197	0.001015
	$\beta_1 = 0.1$	0.001533	0.002062	0.000218	0.04445	0.03411	0.034837
	$\beta_2 = 0.2$	0.001074	0.002144	0.000676	0.0138	0.036448	0.01675
n = 50	$\delta = 0.01$	0.003591	0.002458	0.000248	0.026558	0.029604	0.003962
	$\beta_1 = 0.1$	0.001549	0.001342	0.000189	0.006817	0.066817	0.031366
	$\beta_2 = 0.2$	0.000431	0.000813	0.00047	0.002421	0.035748	0.075507
n = 100	$\delta = 0.01$	0.001157	0.00071	0.001316	0.014477	0.011172	0.004114
	$\beta_1 = 0.1$	0.000583	0.000389	0.000369	0.001246	0.005791	0.023053
	$\beta_2 = 0.2$	0.002605	0.001769	4.23E-05	0.008548	0.086216	0.089184

4.2. Comparing Bootstrap and Bayesian approaches using MSE

Table 4 shows the comparison of the two techniques under consideration in terms of the mean square errors. Mean square error shows the degree of squared deviations of the errors produced by the estimators. The bootstrapping technique produced smaller MSEs than the Bayesian estimator for all sample sizes. The bootstrap method exhibits an asymptotic property, that is, the mean square errors of the estimators decrease consistently as the sample size increases across all number of boots considered.

For small sample sizes, n=20, 30 and 50, least values of MSE are obtained when the number of boots is 1000. However, for large sample (n=100) the values of the MSE are similar for all the boots (boot = 500, 1000 and 5000), see Fig. 1.

For small sample sizes, n=20, 30 and 50, least values of MSE are obtained when the number of boots is 1000. However, for large sample (n=100) the values of the MSE are similar for all the boots (boot = 500, 1000 and 5000), see Figure 1. On the other hand, for the Bayesian method, the minimum MSE values are obtained at 500 replications for all the sample sizes while the highest values of MSE are recorded at 5000 replications, see Fig. 2. Hence, when estimating the parameters of a dynamic model, optimum results can be obtained at 500 replications as better estimates are not likely to be obtained at higher replications.

Table 4
Performance of Bootstrap and Bayesian approaches using MSE

Sample size	Coefficient	Bootstrapping			Bayesian		
		Boot			Replication		
		500	1000	5000	500	1000	5000
n = 20	$\delta = 0.01$	0.031001	0.029618	0.030257	0.147502	0.033106	0.034782
	$\beta_1 = 0.1$	0.005069	0.005093	0.005113	0.026238	0.170715	0.165638
	$\beta_2 = 0.2$	0.006529	0.006594	0.006261	0.023919	0.01795	0.017687
n = 30	$\delta = 0.01$	0.017133	0.016924	0.018138	0.110989	0.007733	0.013975
	$\beta_1 = 0.1$	0.003094	0.00288	0.002964	0.017112	0.021635	0.023792
	$\beta_2 = 0.2$	0.003813	0.003564	0.003439	0.016917	0.015894	0.02005
n = 50	$\delta = 0.01$	0.014676	0.014625	0.015103	0.080529	0.01252	0.003621
	$\beta_1 = 0.1$	0.00205	0.002014	0.002101	0.013465	0.083674	0.07753
	$\beta_2 = 0.2$	0.001917	0.001867	0.001858	0.012201	0.010524	0.016019
n = 100	$\delta = 0.01$	0.009363	0.009789	0.009417	0.006037	0.006009	0.005951
	$\beta_1 = 0.1$	0.001242	0.00127	0.001296	0.050703	0.049009	0.050214
	$\beta_2 = 0.2$	0.00127	0.001376	0.001354	0.006606	0.013681	0.014508

Figure 1
Mean square error of the lagged parameter using Bootstrapping technique
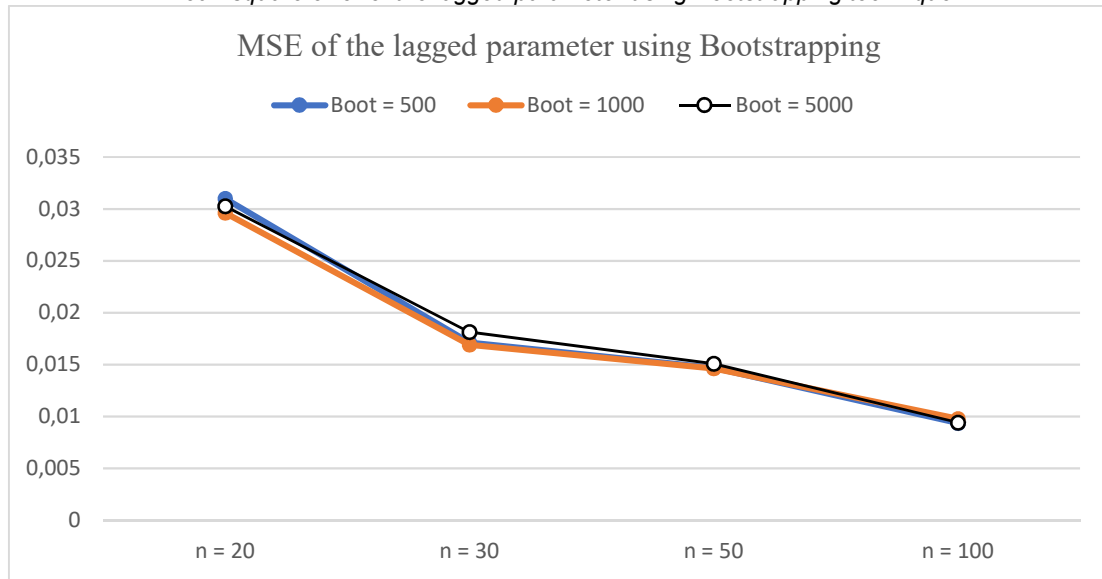
Figure 2

Mean square error of the lagged parameter using Bayesian approach
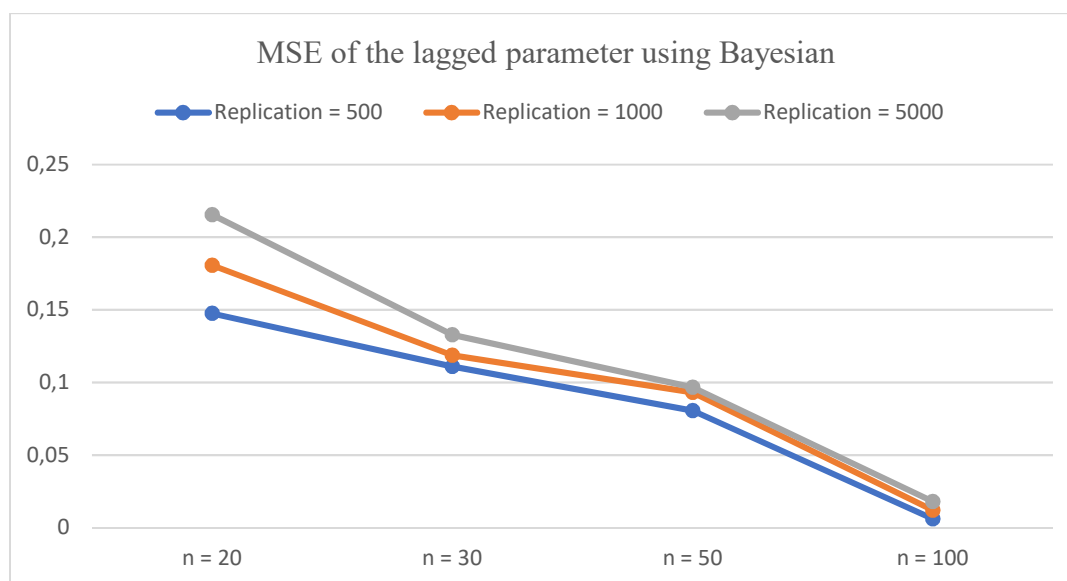
5. CONCLUSION

In this paper, we evaluated bootstrap and Bayesian approaches for estimating the parameters of a dynamic model. The bootstrap technique used was residual resampling approach while for the Bayesian approach, the normal distribution was used for the likelihood and gamma distribution as the conjugate prior.

The simulations showed that bootstrapping only residuals greatly reduce values produced for the standard deviations of parameters.

Another interesting result of this study is the good performance of bootstrap method. The bootstrap method displayed an asymptotic property with the mean square errors of the estimators decreasing as the sample size increases at all boots of 500, 1000 and 5000.

In conclusion, the bootstrap performed better than the Bayesian method when modelling dynamic linear equations. Theoretically, the residual bootstraps always generate datasets with the same design as the original data, it is therefore expected to perform better in situations where the design is not similar for every individual.

REFERENCIAS BIBLIOGRÁFICAS

BERKOWITZ, J. AND KILIAN, L. (2000). Recent Developments in Bootstrapping Time Series, Econometric Reviews, Vol. 19, 1 - 48.

CHERNICK, M. R. (2007). Bootstrap methods: a guide for practitioners and researchers.2nd ed. Hoboken, New Jersey: John Wiley & Sons, Inc.

DAVISON, A.C. AND HINKLEY, D.V. (1997). Bootstrap Methods and their Applications. Cambridge University Press, Cambridge.

EFRON, B. (2012). Bayesian Inference and the Parametric Bootstrap. The Annals of Applied statistics, Vol. 6, No. 4.

EFRON, B. AND TIBSHIRANI, R.J. (1993). An Introduction to the Bootstrap. Chapman and Hall, London.

HALL, P. (1986). On the Number of Bootstrap Simulations Required to Construct a Confidence Interval. The Annals of Statistics. Vol. 14, 1456 – 1462.

HALL, P. (1989). Unusual Properties of Bootstrap Confidence Intervals in Regression Problems. Probability Theory and Related Fields, Vol. 81, 247 – 273.

HALL, P. (1992). The Bootstrap and Edgeworth. Springer Series in Statistics, New York

HANSEN, B. (1999). The Grid Bootstrap and the Autoregressive Model. The Review of Economics and Statistics. Vol. 81, 594 – 607.

HOROWITZ, J.L. (1997). Bootstrap Methods in Econometrics: Theory and Numerical Performance, in: Kreps, D., and Wallis, K., eds., Advance in Economics and Econometrics: Theory and Applications (Cambridge University Press, Cambridge), Vol. 3, 188 – 222.

LI, H. AND MADDALA, G.S. (1996). Bootstrapping Time Series Models (with discussion), Econometric Reviews, Vol. 15, 115 – 158.

MADDALA, G.S. AND JEONG, J. (1993). A perspective on application of bootstrap methods in econometrics, in G.S. Maddala, C.R. Rao, and H.D. Vinod, eds., Handbook of Statistics vol.11 (North-Holland, Amsterdam).

THAI, H.T., MENTRE, F., HOLFORD, N. VEYRAT-FOLLET, C.AND COMETS, E. (2014). A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models. Pharmaceutical Statistics, Wiley, 12 (3), pp.129-40.

ZAHARI, S.M., RAMLI, N.M. AND MOKHTAR, B. (2014). Bootstrapped Parameter Estimation in Ridge Regression with Multicollinearity and Multiple Outliers. J. Appl. Environ. Biol. Sci., 4(7S), 150-156.

ZHANG, L. AND WORTHINGTON, A. C, (2015). Evaluating the accuracy of scope economies: comparisons among delta method, bootstrap, and Bayesian approach. Australian Conference of Economists PhD Colloquiu

Appendix

Figure 3-6 showing the standard deviation histogram for the Bootstrapping Result.

Fig 3
Standard deviation at n=20 for bootstrapping

Fig 4
Standard deviation at n=30 for bootstrapping
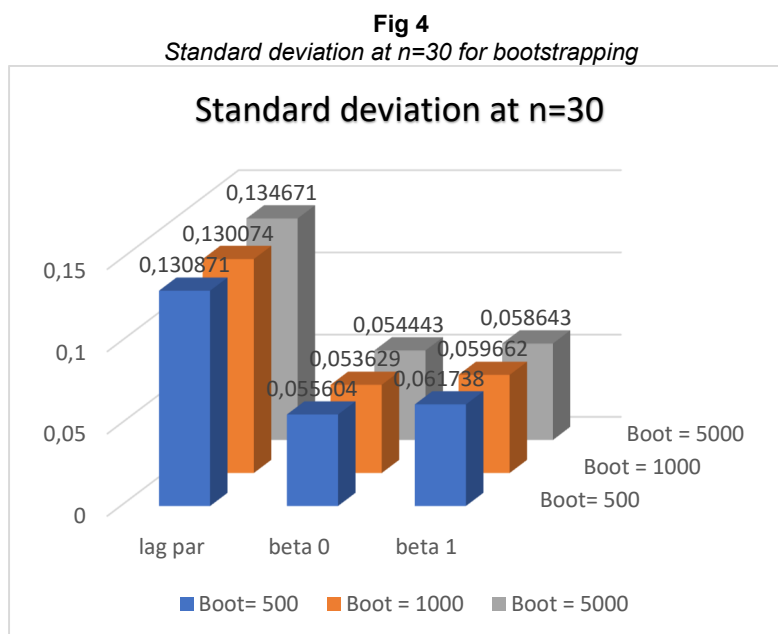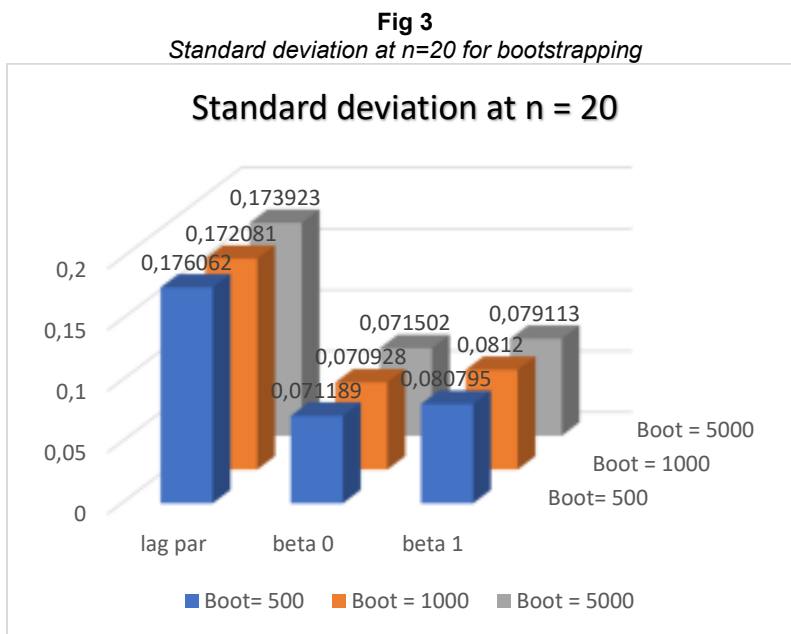
Fig 5
Standard deviation at n=50 for bootstrapping
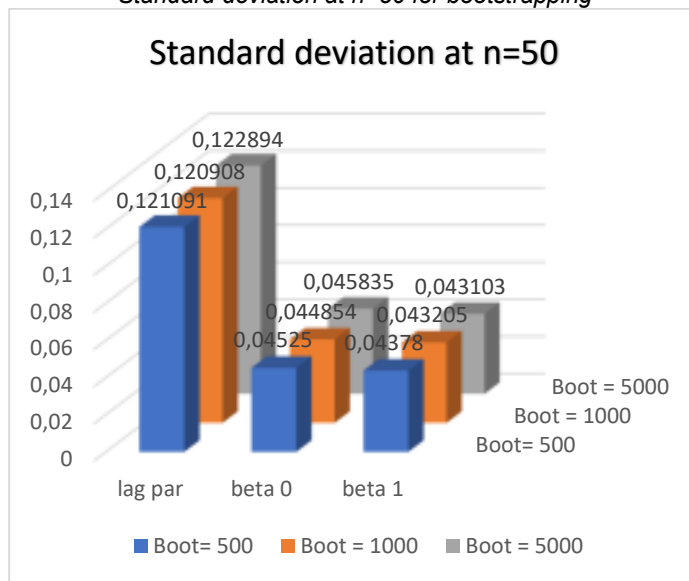
Fig 6
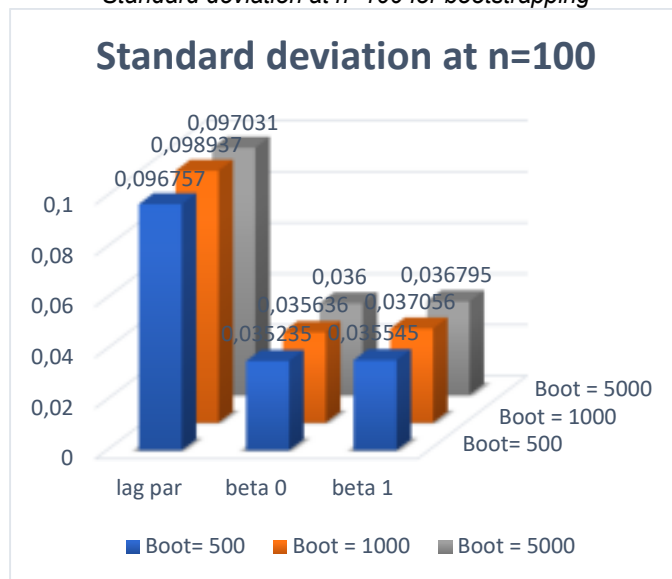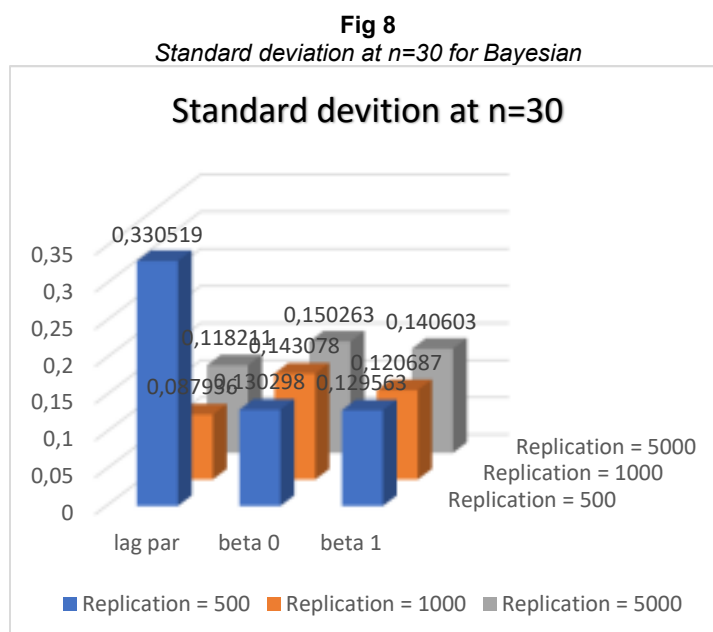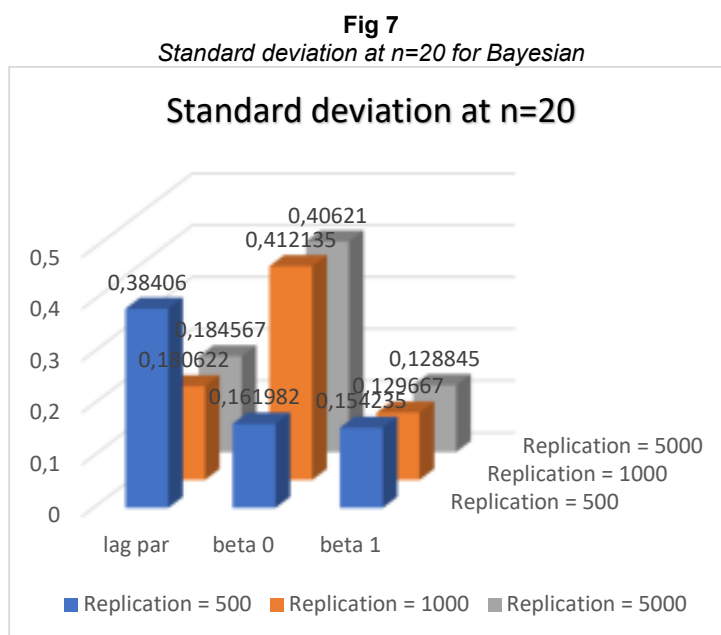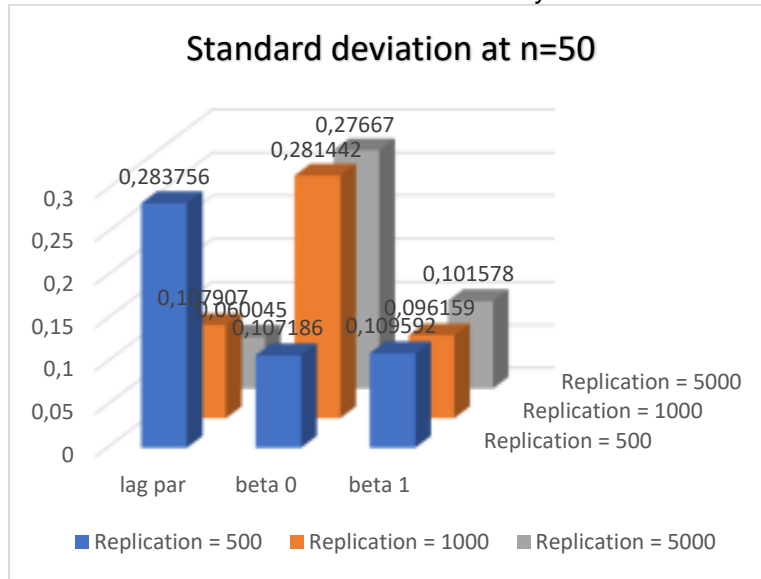*Standard deviation at n=100 for bootstrapping*

Figure 7-10 showing the standard deviation histogram for the Bayesian Result

Fig 7
Standard deviation at n=20 for Bayesian

Fig 8
Standard deviation at n=30 for Bayesian

Fig 9
Standard deviation at n=50 for Bayesian

Fig 10
Standard deviation at n=100 for Bayesian