An Investigation of Bootstrap Methods in Parametric Estimations in Simple Linear Regression

Acha, Chigozie K[1*], Nwabueze, Joy C.[1]

[1] Department of Statistics, Michael Okpara University of Agriculture, Umudike, Abia State, Nigeria

Email Address

acha.kelechi@mouau.edu.ng (Acha, Chigozie K)
*Correspondence: acha.kelechi@mouau.edu.ng

Received: 5 June 2018**; Accepted:** 25 June 2018**; Published:** 17 July 2018

Abstract:

This study compares the bootstrap methods in parametric estimations in simple linear regression. To achieve this set objective theoretical datasets replicated from normal distribution with different ability levels were used. Simple linear regression (SLR) models were employed to fit the datasets and all the scores were considered. Evidence showed that the original data set $(O310M_t)$ without bootstrap resulted in a poor model with high bias. Results also showed that when the sample size was small $\leq$ 200, in almost all the different assessment conditions, the parametric bootstrap model $(H311M_t)$ performed better than all of the parametric bootstrap models including $(O311M_t)$ by showing the smallest conditional bias. The parametric bootstrap model $(H311M_t)$ based on performance is followed by $(H313M_t)$ then $(H312M_t)$ and $(O310M_t)$. Across all the bootstrap conditions, it was obvious that all the models that worked well had the highest information criteria values: HQIC, SBIC, and AIC with standard error term $\leq$ 0.005, minimum bias and RMSE (root mean square error) confirming that the models are good models for further studies and predictions in the economy. Based on the evidence in this study, the paper concludes that parametric bootstrap estimation methods at different assessment conditions can be used to infer the statistical properties of the estimators and constructed inference based on hypothetical data sets.

Keywords:

Parametric Bootstrap, Linear Regression, Estimations, information Criteria, Models

1. Introduction

In statistics, resampling is approximating the accuracy of sample statistics by using subsets of accessible datasets or drawing arbitrarily with replacement from a set of data points. The bootstrap method is a resampling method that reduces error and provides more reliable statistical inference. Bootstrap can be used to make inferences about some experimental results when the statistical theory is not certain or not known. We can also use the bootstrap to assess how well the statistical theory holds: that is, whether an inference we make from a hypothesis test or confidence interval is justified. Bootstrapping in program development began during the 1950s when each program was constructed on paper in decimal code or in binary code, bit by bit (1s and 0s), because there was no high-level computer language, no compiler, no assembler, and no linker as shown in [1].

The importance of bootstrap methods for the purpose of reducing error and providing more reliable statistical inference can never be over emphasized. The simulation study presented in [2], comparatively analyzed weighted least squares (WLS), ordinary least squares (OLS), and three versions of bootstrap (resampling residuals, resampling of data points, generating new residuals from Laplace distributions) for a linear regression with independent residuals from a mixture of two Laplace distributions. Other researches that have worked extensively using different bootstrap methods, [3-18], but to the best of my knowledge, none has compared the bootstrap methods in parametric estimations in simple linear regression by using theoretical sets of data computer-generated with different ability levels from normal distribution.

In spite of all the proposed tests and bootstrap estimation methods in literature, a number of questions still remain unanswered in connection with the proposed test and estimation methods in bootstrap. There is still need to improve on the parametric (PB) estimation methods. Apart, from the objective of investigating the parametric methods, this study is also poised to compare the parametric bootstrap methods in the parameter estimation of the simple linear regression (SLR) under a variety of assessment conditions. The models from this research work will be useful to government in predicting and forecasting trends in the Nigeria economy.

The interest for this study was ignited by the research work presented in [19], which called for more research on the parametric bootstrap method and for comparative studies of bootstrap approaches. Also, the fact that parametric models at the sampling stage of the bootstrap methodology lead to procedures which are different from those obtained by applying basic statistical theory. Therefore, this research work is poised to investigate and understand the parametric bootstrap methods and to compare the parametric bootstrap methods in the parameter estimation of the simple linear regression (SLR) under a variety of assessment conditions: standard errors (precision), bias and root mean square error, the sampling distribution, bootstrap distribution of the sample will be from a normal distribution of different forms. Moreover, other information criteria; AkaikeInformation criterion (AIC), Akaike Final Prediction Error (FPE), Schwart Bayesian Information criterion (SBIC), and Hannan-Quinn Information criterion (HQIC) will be considered. It is pertinent to note that, all these estimation methods and information criteria together define the performance of an estimator.

The basic idea of this study is to investigate and compare the different bootstrap parametric estimations in simple linear regression under a variety of assessment conditions. Also to examine the parametric estimation Bootstrap Models, when the error term are independent and identically distributed.

Specifically, the study will be carried out to;

Examine the Parametric Bootstrap estimation methods in SLR Models

Select the Parametric Bootstrap estimation methods in practice under various assessment conditions.

2. Materials and Methods

There are many techniques to specify the bootstrap for models as simple as the linear regression model. The simplest approach is to estimate the model that corresponds to the null hypothesis and then use the estimates to generate the bootstrap samples, under the assumption that the error terms are normally distributed.

$$\mathrm{yt} = \mathrm{Xt}\,\beta + \mathrm{ut}, \qquad \sim \mathrm{NID}(0,\sigma^2) \tag{1}$$

The model (1) is a fully specified parametric model, which implies that parameter values for each set of β, σ^2 defines just one function. The simplest type of bootstrap for fully specified models is given by the parametric bootstrap. The first step in constructing a parametric bootstrap is to estimate (1) by OLS, yielding the restricted estimates $\tilde{\beta}$, and $\tilde{s}^2$. Then the bootstrap is given by

$$y_t^* = X_t\tilde{\beta} + \mu_t^*, \qquad \mu_t^* \sim \mathrm{NID}(0, \tilde{s}^2) \tag{2}$$

this is the element of the model (1) described by the parameter estimates, with stars to indicate that the data are simulated. To draw a bootstrap sample from the bootstrap (2), we first draw an *n*--vector μ_t^* from the N(0, $\tilde{s}^2$) distribution. For each of the *B* bootstrap samples, θ_k^* , a bootstrap test statistic θ_k^* is computed from y_k^* in just the same way as $\hat{\theta}$ was computed from the original data, *y*.

The residual bootstrap for regression models and their transformations will be discussed, when the data is independently and identically distributed.

(i) The residual bootstrap

One of the approaches to parametric bootstrapping in regression problems is to resample residuals.

Step 1:

Fit the model and retain the fitted values $\hat{y}_i$ and the residuals $\hat{\epsilon}_i = y_i - \hat{y}_i, (i = 1, \ldots, n)$.

Step 2:

For each pair, (x_i, y_i), in which x_i is the (possibly multivariate) explanatory variable, add a randomly resampled residual, $\hat{\epsilon}_j$, to the response variable y_i. In other words create synthetic response variables $y_i^* = \hat{y}_i + \hat{\epsilon}_j$ where *j* is selected randomly from the list (1, …, *n*) for every *i*.

Step 3:

Refit the model using the fictitious response variables y^*_i, and retain the quantities of interest (often the parameters, $\hat{\mu}_i^*$, estimated from the synthetic y^*_i).

Step 4:

Repeat steps 2 and 3 a statistically significant number of times.

Assuming the error terms in (1) are independent and identically distributed with common variance σ^2, then we can infer accurately using the residual bootstrap. Firstly, it is necessary in the residual bootstrap to find OLS estimates β and residuals $\hat{\mu}_t$. It is advisable to rescale the residuals so that they have the correct variance, unless the quantity to be bootstrapped is invariant to the variance of the error terms. The simplest type of rescaled residual is

$$\ddot{\mu}_t \equiv \left(\frac{n}{n-k}\right)^{1/2} \hat{\mu}_t \quad (3)$$

The first factor here is the inverse of the square root of the factor by which 1/n times the sum of squared residuals underestimates σ^2. A more intricate approach uses the diagonals of the 'hat matrix'

$$(XTX)-1XT \quad (4)$$

to rescale each residual by a different factor. Davidson and MacKinnon in their work showed that the rescale approach which uses the diagonals of the hat matrix is a little better than (3) when some observations have high leverage [20]. Though it is known that the goal of leverage in regression analyses is to identify observations that are not near to their corresponding average predictor values. It is pertinent to note that points of leverage may not largely effect the outcome of fitting regression models. Using rescaled residuals the residual bootstrap produces an observation of the bootstrap sample by (5)

$$y_t^* = X_t\hat{\beta} + \mu_t^*, \qquad \mu_t^* \sim EDF(\ddot{\mu}_t) \quad (5)$$

The bootstrap errors in (5) are said to be 'resampled' from the (3). Implying that they are drawn from the empirical distribution function, or EDF, of the $\ddot{u}t$. This function assigns probability 1/n to each of the $\ddot{u}t$ and each of the bootstrap error terms can take on n possible values.

2.1. Evaluation Criteria

The following statistical evaluation criteria were used to investigate and understand the bootstrap methods and to compare the different types of bootstrap for the simple linear regression (SLR) models under different assessment conditions. Generally, to evaluate a satisfactory degree of performance and validity of the methods for the hypothetical data sets and its best model, several assessment conditions were evaluated 200 times to estimate the standard error and bias from the two bootstrap methods in this paper. In bias test, a difference of 0.1 standard deviation units is generally considered relatively large, whereas a difference of 0.25 is regarded as very large, [21] and [19]. This position will be adopted in this paper.

These several assessment conditions are described below

i. Bootstrap method as indicated earlier, the parametric bootstrap (PB) and nonparametric bootstrap (NPB) methods were considered.
ii. The bootstrapped ability levels investigation and evaluation will be described in the same form N(0,1). Here, we use the standard error to get the group differences.
iii. Different sample sizes of three tests lengths: $n_1 = 200$, $n_2 = 1000$ and $n_3 = 10000$ were studied. These levels represented typical small, medium, and large sample sizes.
iv. Bootstrap levels (PB and NPB levels): B-Level will be 99, 499, 1999. These levels also represented typical small, medium, and large sample sizes.
v. Model selection criteria: In this section, several information criteria will be used to select the best model among bootstrap. In-sample will be considered since it essentially tells us how the chosen model fits the data in a given sample. The formulas for other information criteria used to generate the values in Table 3 are shown as follows:

a. Adjusted R^2 criterion

As a penalty for adding regressors to increase the R^2value, Henry Theil developed the adjustedR^2. Formula for adjusted R denoted by $\bar{R}^2$ is

$$\bar{R}^2 = 1 - \frac{RSS/(n-k)}{TSS/(n-1)} = 1 - (1 - R^2)\frac{n-1}{n-k} \quad (6)$$

It is pertinent to note that $\bar{R}^2 \leq$ R2, unlike R2, the adjusted R2 will increase only if the absolute t value of the added variable is greater than 1.

b. Akaike information criterion (AIC):

The idea of imposing a penalty for adding regressors to the model has been carried further in the AIC criterion, which is defined as:

$$AIC = e^{2k/n}\frac{\sum \hat{\mu}_i^2}{n} = e^{2k/n}\frac{RSS}{n} \quad (7)$$

where k is the number of regressors (including the intercept) and n is the number of observations. For mathematical convenience, equation 6, is written as

$$\ln AIC = \left(\frac{2k}{n}\right) + \ln\left(\frac{RSS}{n}\right) \quad (8)$$

where ln AIC = natural log of AIC and 2k/n is penalty factor. In comparing two or more models, the model with the lowest value of AIC is preferred. One advantage of AIC is that it is useful for not only in-sample but also out-of-sample performance of a regression model [22].

c. Schwart bayesian information criterion (SBIC):

The only difference between AIC and SBIC is that SBIC imposes a harsher penalty than AIC [23]. SBIC criterion is defined as:

$$SBIC = n^{k/n}\frac{\sum \hat{u}_i^2}{n} = n^{k/n}\frac{RSS}{n} \quad (9)$$

Equation 7 can also be written as:

$$\ln SBIC = \frac{k}{n}\ln. n + \ln\left(\frac{RSS}{n}\right) \quad (10)$$

where [(k/n) ln. n] is the penalty factor. Like AIC, the lower the value of SBIC, the better the model. Again, like AIC, SBIC can be used to compare in-sample or out-of-sample performance of a model.

d. Hannan-Quinn information criterion (HQIC)

In statistics, the Hannan-Quinn Information Criterion (HQIC) is a criterion for model selection. It is an alternative to Akaike Information Criterion (AIC) and Schwart Bayesian Information Criteria (SBIC) or Bayesian Information Criterion (BIC). It is given as:

$$HQIC = n\log\left(\frac{RSS}{n}\right) + 2k\log\log n \quad (11)$$

where; k is the number of parameters, n is the number of observations, and RSS is the residual sum of squares that results from linear regression or other statistical model [24].

e. Bias in Bootstrapped Regression Models

The algorithm for estimating standard errors from regression models as suggested in [9] will be used,

$$\text{bias} = \text{bias}\left(\hat{\theta}, \theta\right) \quad (12)$$

where; $\hat{\theta} = s(X)$ and $\theta = t(F)$

This means that the bias of $\hat{\theta}(b) = s(x)$ as an estimate of θ is defined to be difference between the expectation of $\hat{\theta}$ and the value of the parameter θ.

3. Results and Discussion

In this study, two groups of models: original and parametric hypothetical data sets were selected to represent the functional parametric models after more than 200 trials were carried out within each bootstrap level. In fact, the parametric models were further divided into four;

a. Ordinary residual method using raw residuals
b. Ordinary residual method using studentized residuals
c. Rescaled residual method using the diagonals of the hat matrix on studentized residuals.
d. Rescaled residual method using the inverse of the square root of the factor on studentized residuals.

Here, the original analysis of the data sets will be carried out. Recall (3),

$$yt = Xt\,\beta + ut, \tag{13}$$

which is now called (13) will be used to estimate original hypothetical sets with fixed sample size is as follows;

i. Ordinary Residual Hypothetical Model $O311M_t$ using raw residuals:

$$O311M_t = f(b1, b2) \tag{14}$$

Original Residual hypothetical Model $O311M_t$: B= 99, N(0,1), n= 30

$$O311M_t = (62.128b1 + 50.662b2) \tag{15}$$

$$\begin{array}{lll} Standard\ error & (0.0019) & (0.0029) \\ Bias & (0.0017) & (0.0025) \\ RMSE & (0.0002) & (0.0041) \\ AIC_{O311M_t} & (90.8269) & \end{array}$$

Equation 15, shows that the model predicts that one unit increase in the b2; the $O311M_t$ will increase by 0.0729 units holding b1 constant. In addition, the model predicts that one unit increase in the b1; the $O311M_t$ will increase by 52.128 units holding b2 constant. The same interpretation goes for equation (17), (19), and (21).

ii. Rescaled Residual Hypothetical Model using the diagonals of the hat matrix on studentized residuals $H311M_t$:

$$H311M_t = f(b1, b2) \tag{16}$$

Rescaled Residual Hypothetical Model $H311M_t$: B= 499, N(0,1), n3=10000

$$PSMC_t = (81.934b1 + 69.025b2) \tag{17}$$

$$\begin{array}{lll} Standard\ error & (0.0005) & (0.0004) \\ Bias & (0.0001) & (0.0015) \\ RMSE & (0.0002) & (0.0031) \\ AIC_{\mathrm{H311M}} & (98.5229) & \end{array}$$

iii. Rescaled Residual Hypothetical Model using the inverse of the square root of the factor on studentized residuals $H312M_t$:

VOLUME 2, 2018
DOI: 10.31058/j.data.2018.22001

$$H312M_t = f(b1, b2) \tag{18}$$

Rescaled Residual Hypothetical Model $H312M_t$: B= 499, N(0,1), n3=10000

$$PSMC_t = (71.937b1 + 64.023b2) \tag{19}$$

$$\begin{array}{lll} Standard\ error & (0.0015) & (0.0012) \\ Bias & (0.0013) & (0.0010) \\ RMSE & (0.0004) & (0.0022) \\ AIC_{\mathrm{H312M}}\ (93.2284) & & \end{array}$$

iv. Ordinary Residual Hypothetical Model using studentized residuals $H313M_t$:

$$H313M_t = f(b1, b2) \tag{20}$$

Ordinary Residual Hypothetical Parametric estimation Model $H313M_t$: B= 499, N(0,1), n3=10000

$$PSMC_t = (64.921b1 + 58.044b2) \tag{21}$$

$$\begin{array}{lll} Standard\ error & (0.0012) & (0.0006) \\ Bias & (0.0005) & (0.0011) \\ RMSE & (0.0010) & (0.0032) \\ AIC_{\mathrm{H313M}}\ (95.4111) & & \end{array}$$

The basis for selection was the fact that as n increases the same pattern is sustained by the models, another model will not be selected when there is a change in pattern. Equations (15), (17), (19) and (21) represent each of the groups of models selected, while results from the analysis are presented in Tables (1- 4).

Table 1. *Bias of the SLR for Parametric estimation Models Comparison.*

Bootstrap Level	Ability Level	Sample Size	Bias for the Bootstrap Models			
			$O311M_t$	$H311M_t$	$H312M_t$	$H313M_t$
B=99	N(0,1)	200	**0.0006**	0.0013	0.0009	**0.0006**
		1000	0.0007	0.0007	**0.0004**	0.0008
		10000	0.0017	**0.0001**	0.0013	0.0005
B=499	N(0,1)	200	0.0024	0.0004	0.0009	**0.0002**
		1000	0.0008	0.0005	**0.0003**	**0.0003**
		10000	0.0019	**0.0001**	0.0018	0.0006
B=1999	N(0,1)	200	0.0004	**0.0002**	0.0006	0.0036
		1000	0.0006	**0.0003**	0.0011	0.0011
		10000	0.0005	**0.0001**	**0.0001**	0.0007

Note. Bold values signify smallest bias in the parametric estimation bootstrap models in each row-Table 1.

Table 2. *Standard error of the SLR for Parametric estimation Models Comparison.*

Bootstrap Level	Ability Level	Sample Size	Standard error for the Bootstrap Models			
			$O311M_t$	$H311M_t$	$H312M_t$	$H313M_t$
B=99	N(0,1)	200	0.0024	0.0002	0.0015	**0.0001**
		1000	0.0036	0.0023	**0.0021**	**0.0021**
		10000	0.0019	**0.0005**	0.0015	0.0012
B=499	N(0,1)	200	0.0016	**0.0007**	0.00I6	0.0082
		1000	0.0075	0.0007	0.0121	**0.0006**
		10000	0.0023	**0.0002**	**0.0002**	0.0003
B=1999	N(0,1)	200	0.0029	**0.0001**	0.0008	**0.0001**
		1000	0.0050	**0.0002**	0.0032	0.0008
		10000	0.0022	**0.0000**	**0.0000**	0.0019

Note. Bold values signify smallest standard error in the parametric estimation bootstrap models in each row-Table 2.

Table 3. *RMSE of the SLR for Parametric estimation Models Comparison.*

Bootstrap Level	Ability Level	Sample Size	Standard error for the Bootstrap Models			
			$O311M_t$	$H311M_t$	$H312M_t$	$H313M_t$
B=99	N(0,1)	200	**0.0002**	0.0008	0.0005	0.0018
		1000	0.0003	**0.0001**	0.0008	**0.0001**
		10000	0.0006	**0.0002**	0.0004	0.0010
B=499	N(0,1)	200	**0.0002**	0.0009	0.00I9	0.0282
		1000	0.0007	**0.0004**	0.0024	0.0020
		10000	0.0007	**0.0005**	0.0006	**0.0005**
B=1999	N(0,1)	200	0.0019	**0.0011**	0.0008	**0.0011**
		1000	0.0016	**0.0002**	0.0020	0.0019
		10000	0.0001	**0.0000**	**0.0000**	0.0002

Note. Bold values signify smallest RMSE in the parametric estimation bootstrap models in each row-Table 3.

Table 4. *Selection of the best model: H311M, H312M, and H313M at the Various Information Criteria.*

Lag	$HQIC_{H311M}$	$SBIC_{H311M}$	AIC_{H311M}	$HQIC_{H312M}$	$SBIC_{H312M}$
0	62.7420	57.5275	72.7112	62.7172	66.5474
1	56.1161	54.3341	57.6321	86.6611	50.4410
2	66.4263	73.6842	77.9909	53.9921	67.0008
3	85.7630	65.0119	69.8974	65.1140	79.9771
4	63.0149	82.5374*	98.5229*	69.7709	88.3320*
5	90.4310*	68.8740	80.0040	91.7234*	67.4000

Table 4. *Cont.*

Lag	AIC_{H312M}	$HQIC_{H313M}$	$SBIC_{H313M}$	AIC_{H313M}
0	54.8852	88.1122	71.2472	79.1240
1	60.1240	64.0110	53.9980	95.4111*
2	85.4111	80.5311	84.5211	41.5323
3	41.5323	69.5540	59.4472	82.7112
4	54.8890	93.3223*	64.7620	57.6321
5	93.2284*	61.5590	92.1861*	77.9909

3.1. Discussion /Interpretation of the Result

In this study, three models were selected to represent the hypothetical bootstrap data sets after more than 200 trials were carried out within each bootstrap level. The basis for selection was that as number of trials (n) increased, the models maintained the same pattern. Another model was selected only when a change in the pattern was noted. The four equations (15), (17), (19) and (21) represent each of the groups of models selected from the Ordinary residual method using raw residuals - $O311M_t$, Rescaled residual method using the diagonals of the hat matrix on studentized residuals - $H311M_t$, Rescaled residual method using the inverse of the square root of the factor on studentized residuals. $H312M_t$, Ordinary residual method using studentized residuals - $H313M_t$ and their results presented in Tables 1, 2, 3, and 4.

The evaluation of the parametric models and their performance will help determine the effects of the factors of sample size, ability level and bootstrap level under different assessment conditions, [25]. Apart from the observation of very low estimates, extreme values in the ranges stated above were reduced. The presentation of the results in these ranges is necessary to show the performance and the trends at the lower ends of the distributions for each bootstrap model. The models provide information to measure the relative effects of sample size, and group proficiency level on the bias, standard error and RMSE.

Tables shows that across all the bootstrap conditions, it was obvious that all the models that worked well had the highest values HQIC, SBIC, and AIC. In order to highlight the differences among the models based on their bias, standard error, RMSE and information criteria, we have $H311M_t$ $better\ than$ $H313M_t$ $followed\ by$ $H312M_t$ $and\ lastly$ $O311M_t$ parametric estimation model under different assessment conditions under this study.

It can be seen from Tables 1 and 2, that sample size and test length of bootstrap level had large effects on bias of the SLR, ability level had relatively small or mixed effects under some conditions bias was smaller for a larger sample size and a shorter test length. Given the same test length, a smaller ratio normally yielded slightly larger bias, especially for the parametric model with the smaller estimate. Table 1 shows that $O311M_t$ and $H313M_t$ models performed very well under all the assessment conditions, at B=99, N(0,1), n=200 with the same standard error value of 0.0006. At this point B=499, N(0,1), n=10000 in Table 3, $H311M_t$ and $H313M_t$ models have the same standard error value – 0.0005. In fact in Tables 1, 2, and 3 the $H311M_t$ and $H312M_t$ have the values – 0.0001, 0.0000 and 0.0000 respectively, with the assessment conditions, at B=1999, N(0,1), n=10000. Although the effect of group proficiency level on bias of the simple linear regression (SLR) was quite small, it seemed there was an interaction effect.

For the $O311M_t$, as the group differences became larger, the bias of the SLR became somewhat smaller; however, for the $H311M_t$ with a longer test, the bias of the SLR became slightly larger as the groups were more different. There was no evidence showing any effect of the group proficiency level on the $H311M_t$ with a short test. Across all the conditions considered, models $O311M_t$model yielded much larger bias than $H311M_t$ model at almost all the estimates. Therefore, for the bootstrap models considered, the pattern was clear that at lower sample sizes and bootstrap levels $O311M_t$ model and other models except $H311M_t$ behave better with respect to bias and standard error. However, it does not mean the higher sample sizes were always associated with the smaller bias along all the estimated values; [4], [5], [11], [8], [9], [10] and [25].

A general observation is that across the group ability level, as the sample size and bootstrap level increased, the bias reduced, meanwhile, the differences among the parametric estimation bootstrap models were becoming more similar. Also, as the sample size, bootstrap level increased, the standard error generally decreased. However, the differences between the results from the parametric estimation models in simple linear regression considered were small (≤ 0.005) and can be recommended for predictions since the determinants are highly significant.

4. Conclusions

When the sample size was small, in almost all the bootstrap conditions, the $H311M_t$ model performed better than all other parametric estimation models by showing the smallest conditional bias, standard error and RMSE. Based on the other information criteria, across all the bootstrap conditions, it was obvious that $H311M_t$, worked well, having the highest values HQIC,SBIC, and AIC, confirming that the models are good models for further studies and predictions in the economic sectors. Based on the evidence in this study, the paper concludes that parametric bootstrap estimation methods at different assessment conditions can be used to infer the statistical properties of the estimators and constructed inference based on hypothetical data sets.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

References

[1] Abney, S. "*Bootstrapping*", ACL '02 Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics Philadelphia, Pennsylvania, USA, 2002, July 07-12.

[2] Alrasheedi, M. Parametric and Non-Parametric Bootstrap: A Simulation Study for a Linear Regression with Residuals from a Mixture of Laplace Distributions. *European Scientific Journal*, 2013, 9(12), 120-131.

[3] Acha, C.K. Impact of Housing and Age of Birds on Egg Production. *International Journal of Numerical Mathematics,* 2010, 5(2), 275-298.

[4] Acha, C.K. Regression and Principal Component Analyses: a Comparison Using Few Regressors. *American Journal of Mathematics and Statistics*, 2012a, 2(1), 1-5, DOI: 10.5923/j.ajms. 20120201.01.

[5] Acha, C.K. On two Methods of Analysis Balanced Incomplete Block Designs. *Pakistan Journal of Statistics and Operation Research*, 2012b, 8(4), 749-757.

[6] Acha, C.K. Parametric Bootstrap Methods for Parameter Estimation in SLR Models. *International Journal of Econometrics and Financial Management,* 2014a, 2(5), 175-179, DOI: 10.12691/ijefm-2-5-2. Available online at http://pubs.sciepub.com/ijefm/2/5/2 (accessed on 23 April 2017).

[7] Acha, C.K. Bootstrapping Normal and Binomial Distributions. *International Journal of Econometrics and Financial Management*, 2014b, 2(6), 253-256, Doi: 10.12691/ijefm-2-6-2.

[8] Cover, T.M.; Thomas, J.A. Elements of Information Theory. John Wiley and Sons. 2006, 254.

[9] Efron, B.; Tibshirani, R.J. An Introduction to the Bootstrap. Chapman and Hall, New York, 1993.

[10] Chernick, M.R. Bootstrap Methods: A Practitioners Guide. Wiley, New York, 1999.

[11] Acha, C.K.; Acha I. A Smooth Bootstrap Methods on External Sector Statistics. *International Journal of Econometrics and Financial Management,* 2015, 3(3), 115-120, DOI: 10.12691/ijefm-3-3-2.

[12] Chernick, M.R.; LaBudde, R. An Introduction to the Bootstrap with Applications in R Wiley, Hoboken, 2011.

[13] Acha, C.K.; Omekara, C.O. Towards Efficiency in the Residual and Parametric Bootstrap Techniques. *American Journal of Theoretical and Applied Statistics,* 2016, 5(5), 285-289, DOI: 10.11648/j.ajtas.20160505.16.

[14] Freedman, D.A. Bootstrapping regression models, Ann. Statist, 1981, 6, 1218-1228.

[15] Lehikoinen, A.; Saurola, P.; Byholm, P.; Linden, A.; Valkama, J. Life history event of euasian sparrow hawk in a changing climate. *Journal of Avian Biology,* 2010, 41, 627-636.

[16] Acha, I.A.; Acha, C.K. Interest Rates in Nigeria: An Analytical Perspective. *Research Journal of Finance and Accounting,* 2011, 2(3), 71-81.

[17] Chatfield, C. The Analysis of Time Series - An Introduction. 6th Edition; Chapman and Hall, London, UK, 2004. ISBN 9781584883173

[18] MacKay, D. Information Theory, Inference and Learning Algorithms. 1st Edition; Cambridge University Press; United Kingdom, 2003, 347-348. ISBN-10: 0521642981; ISBN-13: 978-0521642989.

[19] Wang, T.; Brennan, R.L. A modified frequency estimation equating method for the common-item nonequivalent groups design estimating random error in equipercentile equating. *Applied Psychological Measurement,* 2009, 33, 118-132.

[20] Davidson, R.; MacKinnon, J.G. ‘Bootstrap Methods in Econometrics’, in Patterson, K. and Mills, T.C. (eds), Palgrave Handbook of Econometrics: Volume 1 Theoretical Econometrics. Palgrave Macmillan, Basingstoke, 2006b, 812-38.

[21] Wang, T.; Lee, W.; Brennan, R.L.; Kolen, M.J. A comparison of frequency estimation and chained equipercentile methods under the common-item nonequivalent groups design. *Applied Psychological Measurement,* 2008, 32, 632-651.

[22] Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* 1974, AC-19, 716-723.

[23] Schwarz G. Estimating the Dimension of a Model. *The Annals of Statistics,* 1978, 6(2), 461-464.

[24] Hannan, E.J.; B.G. Quinn. The Determination of the order of an Auto regression. *Journal of the Royal Statistical Society,* 1979, 41, 190-195.

[25] Acha C.K. Rescaling Residual Bootstrap and Wild Bootstrap. *International Journal of Data Science and Analysis,* 2016, 2(1), 7-14, DOI: 10.11648/j.ijdsa.20160201.12.