Comparative Analysis of Technical Efficiency (Bootstrapping Fear Model) among Maize Farmers in Oyo and Osun States of Nigeria

Olapade-Ogunwole, Folayimi

Senior Lecturer, Department of Agricultural Economics & Ladoke Akintola University of Technology, Ogbomoso, Oyo State, Nigeria

Abstracts:-**Agriculture is the mainstay of the Nigeria economy and is characterized by mixed farming system. Expected increases in agriculture require increase in agricultural productivity. Agricultural productivity very much depends on the efficiency of the production process. Policies designed to educate farmers through proper agricultural extension service could have a great impact in increasing the level of efficiency and hence agricultural productivity.**

The bootstrap introduced by Efron (1979) is a method of repeatedly drawing with replacement from a sample. The sample is said to be representative if the moments of its distribution match the moment of the underlying unobservable population. The moment of the bootstrap on the other hand, tend to approach the moments of the observable sample. Therefore, if the sample is a representative one, the bootstrap will yield meaningful results. If it is not representative then the bootstrap will still be consistent in terms of approaching the sample moments, but the results will be counterintuitive.

The need therefore to examine improved maize production technologies vis-à-vis some management practices becomes pertinent in order to identify the factor responsible for the inefficiency of maize farmers. This research therefore focused on the technical efficiency of mazie farmers in Oyo and Osun State Nigeria using fear model.

A multistage sampling technique was used in collecting data from a cross-sectional sample of 300 maize farming households in Oyo and Osun State. Data Envelopment Analysis (DEA) including bootstrapping using FEAR model, Tobit regression analysis were employed in analyzing the data. Results from the Tobit analysis showed that the sigma level for CRS, VRS and scale efficiencies were significant at 1 percent with 24.536, 24.433 and 24.466 respectively.

Keywords: **Maize, Bootstrapping, Oyo and Osun**

I. INTRODUCTION

The bootstrap was first introduced by Efron (1979), while Efron and Tibshirani (1993) provide a nice exposition of various issues associated with bootstrapping. Although it is a well established approach, we need to "re-establish" it for the purposes of this study, emphasizing on certain issues which will help us understand the source of variation and the nature of bias in bootstrap DEA. We will expose our ideas mainly within the regression (OLS) framework, as the principles of bootstrapping within a model are relevant in DEA. A deep understanding on how the bootstrap should be applied on DEA is required in order to design consistent hypotheses to be tested as well as to understand their limitations.

The bootstrap is a procedure of drawing with replacement form a sample, mimicking the data generating process of the underlying true model and producing multiple estimates which can be used for statistical inference. One of its most important uses is to test hypotheses, especially in cases where statistical inference is impossible otherwise. Resampling, within the framework of the bootstrap, relates to redistributing the assumed randomness of the model among observations. This randomness is reflected in the deviation of the model's variable from their expected values, as calculated (or estimated) by the model. The higher the variance of the residuals, the wider the constructed bootstrap confidence intervals will be in hypothesis testing.

In the regression framework (let us assume OLS) these deviations are the model's residuals and there are two methods to bootstrap: to bootstrap pairs (alternatively termed "case resampling") and to bootstrap residuals (or "fixed resampling", as the independent variable is the same in all iteration). In the first case we resample pairs of observations and apply OLS each time. In the second case, we resample residuals by adding them up to the expected value of the dependent variable and apply OLS each time on this new pseudo-variable and the initial independent variables. In each case we obtain a distribution for the estimated coefficient (beta's) of the model which, in the limit, should be equal under both procedures. Resampling residuals is more sensitive to model assumptions (Efron and Tibshirani, 1993), mainly due to the fact that it assumes that the distribution of residual does not depend on the observed sample, it is the same no matter what the independent variable is. However, resampling residuals might be more intuitive and appropriate to be applied in some cases (Efron and Tibshirani, 1993).

The accuracy of the bootstrap estimates depends on two factors: the variance of the model residuals and the inherent bias of the bootstrap process, both of which vary with sample size. Residual variance is the source of variability for bootstrapping and the resulting bootstrap distributions should be similar to the residual distribution (at least the higher moments). In fact, the center of the bootstrap distribution of an estimator is expected to be equal to the valve of the estimator computed by the model. Any deviation from that value is known as the bootstrap bias and it is due to the random resampling process of the bootstrap. Especially if the sample is small and the observations are scattered, the effect of this bias may propagate. Therefore, correcting for bootstrap bias centers the distribution of the estimator to its expected value.

The bootstrap bias should not be confused with the model bias, which is defined as the difference of the model estimates from their true value. The latter occurs when other biases plague the model, which are not always observable, and the two most important ones are the measurement bias and the model specification bias, both of which violate the OLS assumptions in our example. These biases cause the model-estimated parameters to deviate significantly from their "true" value, even asymptotically. Therefore, the bootstrap estimates, which mimic the estimated model, will also fail to converge towards the true values (however, they will still converge towards the model estimates). In fact, considering that bootstrap estimators are also subject to bootstrap bias, it is possible thethey will deviate from the true values even more than the model estimates. Since model biases are unobservable it is impossible to accurately compute the true value of an estimator using the bootstrap distribution; we could only approximate it under the assumption that there are no model biases.

DEA and Bootstrapping

The concept of efficiency has been traditionally related to the ratio of outputs over inputs of a certain firm relative to others. However, in a multiple input-output setup it is necessary to attach weighs to inputs and output, which reflect their relative rate of usage, in order to calculate the ratio of weighted output over weighted input. DEA is a non-parametric technique which is based on this logic and use linear programming to determine optimal weights which minimize the distance between the frontier and the decision making unit (DMU) under consideration, subject to disposability and convexity constraints. The major advantage of DEA is that it does not require the specification of a production function: it just uses a set of inputs the DMUs want to minimize and a set of outputs the DMUs want to maximize.

DEA was first introduced by Charnes, Cooper and Rhodes (1978) with their CRS-consistent "CCR" model, while it as extended by Banker, Charnes and Cooper (1984) to account for VRS. We would like to avoid the exposition of the technical details involved since DEA is well established in the literature. Actually, the intended reader is expected to be already familiar with both DEA and bootstrap DEA methods.

Technical efficiency, as termed in DEA, is most commonly examined under the assumption of either input or output orientation. Under input orientation, DEA efficiency scores are interpreted as required input contractions to make a DMU efficient, keeping the level of output fixed. Under output orientation efficiency scores correspond to required output expansions to make a DMU efficient, keeping input levels fixed. Hence, in input orientation inputs behave as variables and outputs as model parameters. While in output orientation outputs are the variables and inputs the constants.

One of the disadvantages of DEA is that statistical inference is very difficult to be applied on DEA scores. Therefore, bootstrap DEA was introduced by Simar and Wilson (1998), allowing to extract the sensitivity of efficiency scores which results from the distribution of (in) efficiency in the sample. Again, we would like to avoid demonstrating the technical details of the method since it is fairly established, while it would destruct the informed reader from the purpose of the paper. However, further details and analysis on related issues can be found in the papers of Simar and Wilson (1998, 2000) as well as their book chapters (Simar and Wilson, 2004, 2008). The outline of their proposed bootstrap procedure can be summarized in the following steps:

i. Use DEA to calculate efficiency score.
ii. Draw with replacement from the empirical distribution (ED) of efficiency scores. Simar and Wilson (1998) suggest that smoothing the ED provides more consistent results.
iii. Divide the original efficient input level by the pseudo-efficiency scores drawn from the (smoothed) empirical distribution to obtain a bootstrap set of pseudo-inputs.
iv. Apply DEA using the new set of pseudo-input and the same set of outputs and calculate the bootstrapped efficiency scores.
v. Repeat step ii – iv B times and use bootstrapped scores for statistical inference and hypothesis testing.

II. MATERIAL AND METHODS

The study was carried out in Southwestern part of Nigeria precisely Oyo and Osun states. The population of the study comprises all registered maize producing farmers in both states. All agricultural zone in Oyo and Osun States Agricultural Development Projects (OYSADEP) and OSSADEP) were consulted. For administrative convenience, four agricultural zones and thirty three (33) blocks were found in OYSADEP while three agricultural zones and thirty (30) blocks were in OSSADEP.

The Agricultural zones in Oyo states are Ibadan /Ibarapa (14blocks), Ogbomoso (5 blocks) Oyo (5 blocks) and Saki (9 blocks) and those of Osun state include Osogbo (13 blocks/Ife/Ijesha (10 blocks) and Iwo (7blocks). Three agricultural zones were purposively selected from each state making six (6) zones in total, based on the type of crops grown. These were Ogbomoso, Oyo and Saki zone from Oyo state and Ife /Ijesha, Iwo and Osogbo zones from Osun state.

Multistage random sampling technique was employed to sample three hundred (300) maize farmers in the first stage 30 percent blocks were randomly selected from each of the six agricultural zones. A total of sixteen blocks were sampled. Each block comprises eight cells. Second stage, involves random selection of 30 percent of the cell (2) in each block making a total of 32 cells for the study. Finally, 20 percent of the maize farmers in each cell were randomly selected for the study.

Empirical DEA Model

Given that there is an underlying production technology, technical as well as scale efficiencies can be estimated empirically. For a sample of *n* observations of farm households using *k* input to produce *m* outputs the input and output vectors for the *i*th household can be represented as (X_{ki}) and (Y_{mi}) respectively. For a household using (X_{ki}) to produce (Y_{mi}) the input-oriented technical efficiency estimate is defined by:

$$TE(X_{ki}, Y_{mi}) = \underset{\theta,z}{Min\,\theta}\ (\theta, X_{ki}, Y_{mi})$$

$$\text{Subject to} \begin{Bmatrix} y_{mi} \leq \sum_{j=1}^{I} y_{mi}, m = 1, 2, ..., M \\ \sum_{j=1}^{I} Z_i x_{ki} \ \leq \ k = 1, 2, ..., K \\ z_1 \geq 0, i = 1, 2, ... , I \end{Bmatrix}$$

where θ_I = technical efficiency estimate to be calculated for each farm household *i*, y_{mi} = quantity of output *m* produced by farm household *i*, x_{ki} = quantity of input *k* used by farm household *i*, z_i = intensity variable from household *i*, A household is considered to be technically efficient if θ = 1, while a household with θ< 1 is considered to be technically inefficient. The model above assumes constant returns to scale (CRS), which holds that all firms (farm households) operate at the optimum scale (Mugera and Featherstone, 2008). However, because of imperfections in agricultural markets (input/output markets) farms seldom operate under CRS, imposes variable returns to scale (VRS).

$$\sum_{j=1}^{I} Z_i = 1$$

$$\sum_{j=1}^{I} Z_i < 1$$

Tobit Estimates of Determinant of Efficiency of All Farmers in Both (Pooled) States

The overall Tobit determinant of the two states, Oyo and Osun. For CRS, three variables were significant at different level. Age was significant at 1 percent, considering the mean age of the pooled farmers at 47 years. It shows that the farmers in their active ages have a higher rate of adoption of the improved technologies which will enhance efficient production.

Years of farming were significant at 5 percent and negative. Implies that the more experienced that farmers were the less they take the adoption of the technologies into consideration, because they believe in their traditional methods or ways of cultivation.

Household size was significant at 10 percent but negative. It shows that the farm household size has a negative effect on the adoption level of the technologies due to the type of family labour they have.

For scale, four variables were significant which are age, year of schooling, frequency of extension visit and adoption index.

Age was significant at 1 percent and positive. Age has a very significant importance in maize activity and adoption of the technologies. Year of schooling was significant at 1 percent and positive. It implies that the more educated the farmers are, the higher their level of adoption improved technologies.Frequency of extension visit was significant at 1 percent and also positive. It shows that the number of times the extension agent visit the farmers increases in their level of awareness on improve technologies this encourage them to adopt the introduced technologies.Adoption index was significant at 5 percent and positive. In relation to the positive effect of the extension visit, the adoption index of the farmers increases.

The sigma level for the CRS, VRS and scale efficiencies were significant at 1 percent with 24.536, 24.433 and 24.466 respectively, this shows a good fit result.

Tobit Estimates of Determinant of Efficiency of All Farmers in Pooled State

Variable	CRS		VRS		Scale efficiency	
	Coefficient	t-value	Coefficient	t-value	Coefficient	t-value
Constant	0.2440	3.609***	14.446	– 1.757*	0.4287	5.319***
Age	0.4928 – 02	3.206***	0.4313	2.308**	0.5365– 02	2.929***
Years of Schooling	– 0.2580 – 02	– 0.960	0.2559	0.782	0.91897 – 02	2.867***
Frequency of extension visit	– 0.7015 – 02	– 0.584	0.7525	0.515	0.3719 – 01	2.596***
Years of experience	– 0.3211 – 02	–2.290**	–0.2271	– 1.336	– 0.7784 – 03	–0.466
Household size	– 0.8196 – 02	–1.819*	–0.4463	– 0.817	– 0.7428 – 02	–1.383
Adoption	0.1598 – 01	1.487	–0.7811	– 0.598	0.2543 – 01	1.979
Sigma	**0.2049**	**24.536*** **	**24.846**	**24.453*** **	**0.2442**	**24.466*** **

*****1% level of significance, **5% level of significance and *10% level of significance**

Source: Computed from Field Survey, 2017.

Characteristics of Farms With Respect To Returns To Scale

This table revealed the mean farm size (ha) and the mean output (kg) of classified farmers into sub-optimal, optimal and super-optimal. Sub-optimal has 274 farms and mean size of 2.30ha with mean output of 1858.48.

The mean output of super-optimal scale is larger than that of sub-optimal as well as optimal scale for both states. The result indicates that the super-optimal level overlap a substantial portion of sub-optimal and optimal output. This is contrary to Ogunniyi et al (2011)which the work shows a larger optimal output scale over others.

Characteristics of Farms With Respect To Returns To Scale

	Number Of Farms	Mean Farm Size (Ha)	Mea Output (Kg)
Sub-optima (IRS)	274	2.03	1858.48
Optimal (CRS)	12	2.09	4643.33
Super-optimal (DRS)	14	6.25	6342.86

Source: Computed from Field Survey, 2017.

Bootstrapping of Farmers Technical Efficiency

The biasedness for the technical efficiency of variable return to scale, non increase return to scale and constant return to scale comparing the means of technical efficiency and biased was tested. Mean value or VRS was 56.7 while the biased VRS was 50. Means of technical efficiency of VRS was higher to the biased means and 43 farmers (14.3%) were 100 percent efficient in TEVRS while no farmer was fully technically efficient in the biased VRS.

For the mean of non increasing returns to scale and it biasedness, the mean for TEnirs was 34 while biasedness was 27.31, also 14 farmers (4.7%) were technically efficient while 70 farmers was technically efficient for biased Nirs for CRS, the mean for TeCrs was 33.3 while for the biased CRS was 28.2 and it was deduced that 11 farmers (3.7 percent) were fully technically efficient while non was fully technically efficient for its biasedness.

The supper boundary for URS has a higher mean of 55.6 to the lower boundary means of 42.

As a rule of thumb it I recommended that the original technical efficiency estimates should not be corrected for bias unless the absolute value of the estimated bias is greater than a quarter of the estimated standard deviation of the parameter estimated (Efron and Tibshirani, 1993).

Bootstrapping of Farmers Technical Efficiency

Efficiency score	TE VRS	Percen tage	TE NIRS	Percen tage	TE CRS	Percen tage	BCTE VERS	Perce ntage	BCTE NIRS	Percent age	BCTE CRS	Percen tage	UBV RS	Percen tage	LB VRS	Perce ntage
0 – 9	0	0		7	21	7	0	0	30	10	30	10	0	0	0	**0**
10 – 19	0	0	21	16	48	16	10	3.3	81	27	74	24.7	0	0	12	**04**
20 – 29	22	7.3	48	31.3	94	31.3	42	14	95	31.7	94	31.3	26	8.7	84	**28**
30 – 39	78	26	62	20.7	64	21.3	69	23	46	15.3	50	16.7	78	26	60	**20**
40 – 49	45	15	27	9	28	9.3	53	17.7	18	06	24	08	44	14.7	60	**20**
50 – 59	57	19	16	5.3	16	5.3	45	15	13	4.3	08	2.7	56	18.7	35	**11.7**
60 – 69	23	7.7	05	1.7	06	02	20	6.7	07	2.3	05	1.7	19	6.3	21	**07**
70 – 79	08	2.7	04	1.3	05	1.7	21	0.7	10	3.3	07	2.3	12	4	21	**07**
80 – 89	12	4	07	0.7	05	17	28	9.3	0	0	08	2.7	07	2.3	07	**2.3**
90 – 99	12	4	02	4.7	02	0.7	12	04	0	0	0	0	56	18.7	0	**0**
100	43	14.3	14		11	3.7	0	0	0	0	0	0	0	0	0	**0**
Total	300	100	300		300	100	300	100	300	100	300	100	300	100	300	**100**
Means	56.7		34		33.3		50		27.3		28.2		55.6		42	
Maximum	100		100		100		93		79		100		95		88	
Minimum	**25**		**1**		**1**		**1**		**1**		**29**		**29**		**19**	

Source: Computed from Field Survey, 2017.

NB: TE_{VRS} – Technical efficiency of variable return to scale

TE_{CRS} – Technical efficiency of constant return to scale

TE_{NIRS} - Technical efficiency of non increase return to scale

$BCTE_{VRS}$ – Biasedness of technical efficiency of variable return to scale

$BCTE_{CRS}$ – Biasedness of technical efficiency of constant return to scale

$BCTE_{NIRS}$ – Biasedness of technical efficiency of non increase return to scale

UB_{VRS} – Upper boundary of variable return to scale

LB_{VRS} – Lower boundary of variable return to scale

REFERENCES

[1]. Charnes, A., Cooper W.W & Rhodes, E. (1978). "Measuring the inefficiency of decision making units" *European Journal of Operational Research*, 2, : 429 – 444.

[2]. Eron, B. &Tibshirani, R.J. (1993).*An Introduction to the Bootstrap* (New York, Chapman and Hall).

[3]. Mugera, A. W., & Featherstone, A.M. (2008).Backyard Hog Production Efficiency.Evidence from the Philippines.*Asian Economic Journal*, 22(3), 267 – 287. http://dx.dio.org/ 10.1111/j.1467-8381.2008.00278.x.

[4]. Simar L., Wilson W.P., (1998). "Sensitivity Analysis of Efficiency Score: How to Bootstrap in Nonparametric Frontier Models". *Management Science* 44(1):49 – 61.

[5]. Simar L., Wilson W.P., (2000)." Statistical Inference in Nonparametric Frontier Models: the State of Art", *Journal of Productivity Analysis*, 1349 – 78.

[6]. Simar L., Wilson W.P., (2004). Performance of the Bootstrap for DEA Estimators and Iterating the Principle, ed by Copper W.W., Seiford M.L., Zhu J., in Handbook on Data Envelopment Analysis, Kluwer Academic Publisher, pp. 265 – 298.