Applied Studies in Agribusiness and Commerce – APSTRACT
Center-Print Publishing House, Debrecen
DOI: 10.19041/APSTRACT/2016/4-5/16

SCIENTIFIC PAPER

COMPARING OLS AND RANK-BASED ESTIMATION TECHNIQUES FOR PRODUCTION ANALYSIS: AN APPLICATION TO GHANAIAN MAIZE FARMS.

Henry De-Graft Acquah

Department of Agricultural Economics and Extension
University of Cape Coast, Cape Coast, Ghana
e-mail: *henrydegraftacquah@yahoo.com*

Abstract: *This paper introduces the rank-based estimation method to modelling the Cobb-Douglas production function as an alternative to the least squares approach. The intent is to demonstrate how a nonparametric regression based on a rank-based estimator can be used to estimate a Cobb-Douglas production function using data on maize production from Ghana. The nonparametric results are compared to common parametric specification using the ordinary least squares regression. Results of the study indicate that the estimated coefficients of the Cobb-Douglas Model using the Least squares method and the rank-based regression analysis are similar. Findings indicated that in both estimation techniques, land and Equipment had a significant and positive influence on output whilst agrochemicals had a significantly negative effect on output. Additionally, seeds which also had a negative influence on output was found to be significant in the robust rank-based estimation, but insignificant in the ordinary least square estimation. Both the least squares and rank-based regression suggest that the farmers were operating at an increasing returns to scale. In effect this paper demonstrate the usefulness of the rank-based estimation in production analysis.*

Keywords: *Production function, parametric and non-parametric regression, rank-based estimation, ordinary least squares estimation*
(JEL CODE: Q18, D24, Q12, C1 and C67)

INTRODUCTION

Cobb and Douglas (1928) propose an econometric methodology to investigate production functions. This entails specifying a linear relationship between inputs and outputs and estimating the linear model using ordinary least squares estimation technique. Consequently, the parametric estimation of the production function has dominated the literature. However, the Cobb-Douglas econometric technique comes with associated constraints imposed on the data.

Some studies highlight the limitations of the parametric approaches and propose a non-parametric estimation of the production functions. For example Henningsen and Kumbharkar (2009) advertised a semi parametric approach to efficiency analysis that estimates production function by a non-parametric regression approach. Furthermore, some studies (Czekaj and Henningsen, 2011) suggest the use of a non-parametric method to scrutinize the traditional parametric estimation method. Subsequently they provide comparison of parametric and non-parametric estimates of the production function. However these studies proposing a non-parametric estimation do not consider the rank-based non parametric estimation technique. This study expands on the parametric and non-parametric estimation of production functions by exploring the rank based estimation. Rank-based estimators have been developed as robust non parametric alternative to traditional least squares estimators. Rank-based regression was first introduced by Jureckova (1971) and Jaeckel (1972). Mckean and Hettmansperger (1978) developed a Newton step algorithm that led to feasible computation of these rank-based estimates. Kloke and Mckean (2015) developed a package (Rfit) for rank-based estimation and inference for linear models using R programming language. This paper demonstrates that the rank-based non-parametric regression offers an alternative and useful approach to estimating the production function. The paper is outlined as follows. The introduction is followed by the methodology which discusses Cobb-Douglas Production Function, Parametric and non-parametric regression approaches, Ordinary Least Squares and Rank-Based Estimations, Returns to Scale, Results and Discussion, and Conclusion.

METHODOLOGY

The methodology describes the data and the parametric and non-parametric econometric techniques employed in the study. Econometric techniques such as ordinary least squares and rank-based non-parametric regression analysis and the Cobb-Douglas model are emphasized.

Cobb-Douglas Production Function

The Cobb-Douglas function is most commonly used in applied production economics. The Cobb-Douglas production function with *N* inputs is defined as:

$$y = A\prod_{i=1}^{N} x_i^{\alpha_i} \qquad [1]$$

This function can be linearized by taking the (natural) logarithm on both sides:

$$\ln y = \alpha_0 + \sum_{i=1}^{N} \alpha_i \ln x_i \qquad [2]$$

where α_0 is equal to $\ln A$.

Thus, the Cobb-Douglas production function is a linear model of the natural logarithm of both the dependent variable and the independent variable(s). In this study, estimation of the parameters of linearized Cobb-Douglas production function is done using Ordinary Least Squares (parametric) method and Rank-Based estimation (non-parametric), and the results were compared.

Parametric and Non-parametric Regression Approaches

The goal of regression analysis is to estimate the relationship of one or more explanatory variables with a single dependent variable. This is done by evaluating the conditional expectation of the dependent variable given the explanatory variables, which can be expressed as:

$$y_i = f_0(x_i) + \varepsilon_i \qquad [3]$$

$$i = 1, 2, \cdots, n,$$

where $i = 1, 2, \cdots, n$ denotes an observation of a subject, y_i is the response variable, and x_i is a $k \times 1$ vector of predictor variables, $f_0(x_i)$ is the expectation of y_i conditional on x_i (the unknown regression function), and ε_i is the error term.

The traditional parametric approach to regression analysis is to assume that $f_0(x_i)$ belongs to a parametric family of functions: $f_0(x_i|\beta)$. So $f_0(x_i)$ is known to have up to a finite number of parameters. Most importantly, parametric approach to regression analysis requires the specification of a functional form for $f_0(x_i)$.

In non-parametric approach we do not assume a certain parametric functional form for $f_0(x_i)$ but is constructed according to information derived from the data. Nonparametric regression requires larger sample sizes than regression based on parametric models because the data must supply the model structure as well as the model estimates.

Ordinary Least Squares Estimation (OLS)

In this approach, the most crucial decision is the specification of the functional form for $f_0(x_i)$. It is assumed that y_i in the model is linearly related with x_i, and ε_i is independent and identically distributed (iid) with $E(\varepsilon_i) = 0$ and variance σ^2. Consider the following model:

$$f_0(x_i|\beta) = \beta_0 + x_i'\beta \qquad [4]$$

Thus a linear regression model is written as:

$$y_i = \beta_0 + x_i'\beta + \varepsilon_i \qquad [5]$$

where β_0 is the intercept and β is $k \times 1$ vector of parameters. For convenience, Equation 5 can be written as: $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\beta}$ is $p \times 1$ vector of parameters, $p = k + 1$, $\boldsymbol{y}$, is $n \times 1$, $\boldsymbol{X}$ is $k \times p$ design matrix, and $\boldsymbol{\varepsilon}$ is the $n \times 1$ vector of error terms.

Under Gauss-Markov assumptions, the estimators of $\boldsymbol{\beta}$ is the Best Linear Unbiased Estimators (BLUE), and can be estimated by using OLS. Using the OLS method of estimation $\boldsymbol{\beta}$ can be estimated by $\widehat{\boldsymbol{\beta}}$ which is given by the explicit formula:

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \qquad [6]$$

The matrix $(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$ is called the Moore-Penrose pseudo inverse matrix of $\boldsymbol{X}$. After $\boldsymbol{\beta}$ has been estimated, the fitted values (or predicted values) from the regression will be:

$$\widehat{\boldsymbol{y}} = \boldsymbol{X}\widehat{\boldsymbol{\beta}} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} \qquad [7]$$

In the case of simple linear regression (one predictor variable) the model is written as:

$$y_i = \alpha + \beta x_i + \varepsilon_i, \qquad [8]$$

and α and β are estimated as:

$$\hat{\beta} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2} \qquad [9]$$

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \qquad [10]$$

After we have estimated α and β the fitted values (or predicted values) from the regression will be:

$$\hat{y}_i = \hat{\alpha} + \hat{\beta}x_i \quad [11]$$

Notably, the OLS estimator is the minimizer of Euclidean distance between $\boldsymbol{y}$ and $\hat{\boldsymbol{y}} = \boldsymbol{X}\hat{\boldsymbol{\beta}}$.

It is assumed that the errors are independent and identically distributed (iid) with mean 0 and variance σ^2, thus $\boldsymbol{\varepsilon} \sim N(0, \sigma^2\boldsymbol{I})$. Now since $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$, implies that $\boldsymbol{y} \sim N(\boldsymbol{X\beta}, \sigma^2\boldsymbol{I})$, which is a compact description of the regression model. From this it can be found, using the fact that linear combinations of normally distributed values are also normal, that:

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, (\boldsymbol{X}^T\boldsymbol{X})^{-1}\sigma^2)$$

Inference on all the predictor variables can be tested by testing:

$$H_0: \beta_1 = \beta_2 \cdots = \beta_k = 0$$

H_0 is to be rejected if

$$F = \frac{(TSS - RSS)/(p-1)}{RSS/(n-p)} > F_{1-\alpha, p-1, n-p}$$

where $TSS = (\boldsymbol{y} - \overline{\boldsymbol{y}})^T(\boldsymbol{y} - \overline{\boldsymbol{y}})$ which is sometimes known as sum of squares corrected for the mean, and $RSS = (\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})^T(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$ which is the residual sum of squares.

The approximate $(1-\alpha) \times 100\%$ confidence interval for β_j is

$$\hat{\beta}_j \pm t_{1-\alpha/2, n-p}\text{se}(\hat{\beta}_j)$$

where $\text{se}(\hat{\beta}_j) = \hat{\sigma}\sqrt{(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{jj}}$, and $(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{jj}$ is the jth diagonal element of $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$.

Rank-Based Estimation

In contrast, the non-parametric approach to regression analysis does not require any presumptions for the functional form of $f_0(x_i)$. As with OLS, the goal of rank-based estimation is to estimate the vector of parameters, β, of a linear model in Equation 5. For convenience, Equation 5 can be written in matrix notation as:

$$\boldsymbol{y} = \alpha\boldsymbol{1} + \boldsymbol{X\beta} + \boldsymbol{\varepsilon} \quad [12]$$

where $\boldsymbol{y}$ is the $n \times 1$ vector of responses, $\boldsymbol{X}$ is the $n \times k$ design matrix, $\boldsymbol{\beta}$ is $k \times 1$ vector of parameters, and $\boldsymbol{\varepsilon}$ is the $n \times 1$ vector of error terms. The only assumption on the error term is that it is continuous; in that sense the model is general. The geometry of the rank-based procedures is the same as OLS, except that instead of the Euclidean distance, the Jaeckel's dispersion function is used which is based on a pseudo-norm $\|\cdot\|_\varphi$. The Jaeckel's dispersion function is given by:

$$D(\boldsymbol{\beta}) = \|\boldsymbol{y} - \boldsymbol{X\beta}\|_\varphi \quad [13]$$

where $\|\cdot\|_\varphi$ is a pseudo-norm defined as:

$$\|\boldsymbol{u}\|_\varphi = \sum_{i=1}^{n} a\big(R(u_i)\big)u_i \,, \qquad u \in R^n,$$

where the scores are generated as $a(i) = \varphi\left(\frac{i}{n+1}\right)$ for a non-decreasing square-integrable function $\varphi(u)$, defined on the interval $(0, 1)$, and $R(u_i)$ is the rank. Assume without loss of generality that it is standardized, so that $\int \varphi(u)du = 0$ and $\int \varphi^2(u)du = 1$. Two of the most popular score functions are the Wilcoxon $\big(\varphi(u) = \sqrt{12}[u - (1/2)]\big)$ and the L_1 $(\varphi(u) = \text{sgn}[u - (1/2)])$. Because the scores sum to zero and the ranks are invariant to a constant shift, the intercept cannot be estimated using the norm. Instead it is usually estimated as the median of the residuals. That is, $\hat{\alpha}_s^c = \text{med}\{Y_i - x_i^T\hat{\beta}_\varphi\}$, where x_i^T is the ith row of $\boldsymbol{X}$.

The rank-based estimator of $\boldsymbol{\beta}$ is defined as:

$$\hat{\boldsymbol{\beta}}_\varphi = \text{Argmin}\|\boldsymbol{y} - \boldsymbol{X\beta}\|_\varphi \quad [14]$$

This estimator is a highly efficient estimator which is robust in the Y-space. A weighted version can attain 50% breakdown in the X-space at the expense of a loss in efficiency; see Chang et al. (1999).

$\hat{\boldsymbol{\beta}}_\varphi$ is the Hodges-Lehmann estimate (i.e., the median of all pairwise differences between the samples) if the Wilcoxon scores is used. Let $f(t)$ denote the probability density function of ε_i. Then, under regularity conditions:

$$\begin{pmatrix}\hat{\alpha}_s \\ \hat{\boldsymbol{\beta}}_\varphi\end{pmatrix} \text{ is approximately}$$

$$N_{k+1}\left(\begin{pmatrix}\alpha \\ \boldsymbol{\beta}\end{pmatrix}, \begin{bmatrix} k_n & -\tau_\varphi^2\bar{x}^T(\boldsymbol{X}^T\boldsymbol{X})^{-1} \\ \tau_\varphi^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}\bar{x} & \tau_\varphi^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}\end{bmatrix}\right)$$

where $\hat{\alpha}_s = \hat{\alpha}_s^c - \bar{x}^T\hat{\boldsymbol{\beta}}_\varphi$, $k_n = n^{-1}\tau_s^2 + \tau_\varphi^2\bar{x}^T(\boldsymbol{X}^T\boldsymbol{X})^{-1}\bar{x}$, $\tau_s = [2f(0)]^{-1}$,

$\tau_\varphi = [\int \varphi(u)\varphi_f(u)du]^{-1}$, and $\varphi_f(u) = -f'(F^{-1}(u))/f(F^{-1}(u))$.

Depending on knowledge of the error probability density function $f(t)$, appropriate scores can result in asymptotically efficient estimates. This result can be summarize as follows:

$$\hat{\boldsymbol{\beta}}_\varphi \sim N\big(\boldsymbol{\beta}, \tau_\varphi^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}\big)$$

An estimate of τ_φ is necessary to conduct inference. Denote this estimator by $\hat{\tau}_\varphi$. Then Wald tests and confidence intervals can be calculated. Let $\text{se}(\hat{\beta}_j) = \hat{\tau}_\varphi (\boldsymbol{X}^T\boldsymbol{X})^{-1}_{jj}$, where $(\boldsymbol{X}^T\boldsymbol{X})^{-1}_{jj}$ is the jth diagonal element of $(\boldsymbol{X}^T\boldsymbol{X})^{-1}$. Then an approximate $(1-\alpha)\times 100\%$ confidence interval for β_j is

$$\hat{\beta}_j \pm t_{1-\alpha/2,n-p-1}\text{se}(\hat{\beta}_j).$$

A Wald test of the general linear hypothesis

$$H_0: \boldsymbol{M\beta} = \boldsymbol{0} \text{ versus}$$

$$H_1: \boldsymbol{M\beta} \neq \boldsymbol{0}$$

is to reject H_0 if

$$\frac{(\boldsymbol{M}\hat{\boldsymbol{\beta}}_\varphi)^T[\boldsymbol{M}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\boldsymbol{M}^T]^{-1}(\boldsymbol{M}\hat{\boldsymbol{\beta}})/q}{\tau_\varphi^2} > F_{1-\alpha,q,n-p-1}, \text{ where } q = \dim(\boldsymbol{M}).$$

Returns to Scale (RTS)

From the Cobb-Douglas production function, the output elasticities with respect to the factors of production (inputs) are equal to the corresponding coefficients of the Cobb-Douglas regression model. Based on the farmers' output elasticities, it would be known whether the farmers' exhibits constant returns to scale, decreasing returns to scale or increasing returns to scale and its implication to the farmers. The returns to scale is the summation of all the output elasticities of the factors of production. It is specified mathematically as:

$$RTS = \sum_{i=1}^{k} \epsilon_i = \sum_{i=1}^{k} \beta_i \qquad [15]$$

where ϵ_i is the output elasticities with respect to the ith input, and β_i is the coefficient of the ith input of the Cobb-Douglas regression model.

SAMPLE SIZE AND DATA ANALYSIS

In this study, simple random sampling technique was used to select 306 maize farmers from the Ejura Sekyedumase District. The analytical tools used for this study were descriptive statistics and parametric and non-parametric regression analyses. The dependent variable of the production function is the farm's output measured as total maize yield. The independent variables used in the regression analyses were six: labour, land, equipment, agrochemical, fertilizer, and seeds. The R programming software was used to analyse the data. The Cobb-Douglas production function was estimated using the ordinary least squares estimation and the rank-based estimation. The R packages Rfit was used for the rank-based estimation.

RESULTS

Descriptive statistics of the regression variables are presented in Table 1. Findings from Table 1 indicate that on average a yield of 7396.37kg was obtained. This output was obtained by combining 170.65 person-days of labour, 16.06 acres of land, 15.82 litres of agrochemicals, 140.98 kilogram of fertiliser, 5.03 kilogram of seeds and GHS15.68 of equipment.

Table 1: Descriptive Statistics of Regression Variables

Variable	Unit	Minimum	Maximum	Mean	Std. Dev
Output	Kg	480.00	52200.00	7396.37	6919.31
Labour	P-D	28.00	469.00	170.65	75.91
Land	Acres	2.00	60.00	16.06	10.60
Equipment	GHS	2.40	72.00	15.68	14.04
Agrochemicals	Lit.	3.00	63.00	15.82	10.65
Fertiliser	Kg	25.00	300.00	140.98	43.33
Seed	Kg	3.00	9.00	5.03	1.12

OLS estimation results of the Cobb-Douglas regression model presented in Table 2, reveals a significant and positive relationship between land and equipment as explanatory variables and maize yield as the dependent variable. There is also a significant but negative relationship between the use of agrochemicals (weedicides, pesticides, fungicide and insecticide) as an explanatory variable and maize yield as dependent variable. There is also a negative relationship between seed as explanatory variable and maize yield as dependent variable. However, this relationship is not significant.

Table 2: Ordinary Least Square Estimates

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.25876	0.32104	16.380	2e-16 ***
log(Labour)	0.05309	0.04625	1.148	0.25194
log(Land)	1.25648	0.06183	20.321	2e-16 ***
log(Equipment)	0.06933	0.02410	2.876	0.00431 **
log(Agrochemicals)	-0.13983	0.06493	-2.154	0.03207 *
log(Fertilizer)	0.05092	0.05449	0.935	0.35076
log(Seed)	-0.15117	0.07853	-1.925	0.05518
F-test	**Sig.**			
319.3	2.2e-16***			
R-squared				
Multiple R-squared	0.865			
Adjusted R-squared	0.8623			

Sig. codes: ''p< 0.001, '**'p< 0.01, '*'p< 0.05***

Similarly, rank-based estimation results of the Cobb-Douglas regression model presented in Table 3, shows a significant and positive relationship between land and equipment as explanatory variables and maize yield as the dependent variable. Additionally, there is also a significant but negative relationship between the use of agrochemicals (weedicides, pesticides, fungicide and insecticide) and seed as explanatory variables and maize yield as dependent variable.

Table 3: Rank-Based Estimates

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	5.308942	0.331579	16.0111	2e-16 ***
log(Labour)	0.052293	0.047737	1.0954	0.27421
log(Land)	1.271812	0.063816	19.9292	2e-16 ***
log(Equipment)	0.060810	0.024877	2.4444	0.01509 *
log(Agrochemicals)	-0.136052	0.067015	-2.0302	0.04322 *
log(Fertilizer)	0.049349	0.056235	0.8776	0.38089
log(Seed)	-0.181732	0.081055	-2.2421	0.02569 *
Reduction in Dispersion Test		**Sig.**		
163.696		0.000***		
R-squared				
Multiple R-squared (Robust)	0.7666207			

Sig. codes: ''p< 0.001, '*'p< 0.05***

Table 4 presents the elasticities and returns to scale for the OLS estimation and the Rank-based estimation. On the basis of the rank-based Cobb Douglas model, inputs used in producing maize (i.e. labour, equipment, agrochemical, fertilizer and seed) were all inelastic with the exception of land. This was also the case of the OLS estimation.

Table 4: Elasticity of Production and Returns to Scale (RTS)

	OLS Estimation		**Rank-Based Estimation**	
Variable	**Elasticity**	**RTS**	**Elasticity**	**RTS**
Labour	0.05309	1.14	0.052293	1.12
Land	1.25648		1.271812	
Equipment	0.06933		0.060810	
Agrochemical	-0.13983		-0.136052	
Fertilizer	0.05092		0.049349	
Seed	-0.15117		-0.181732	

DISCUSSION

The maximum and minimum yield obtained in Table 1 indicates that there is a large variation in maize output among farmers in the District. The wide variation in output could be attributed to differences in technical efficiency levels of farmers.

In the OLS estimation results of the Cobb-Douglas regression model presented in Table 2, the significant and positive relationship between land and equipment as explanatory variables and maize yield as the dependent variable suggests that an increase in each of these explanatory variables will lead to an increase in the output of maize. The significant but negative relationship between the use of agrochemicals (weedicides, pesticides, fungicide and insecticide) as an explanatory variable and maize yield as dependent variable suggests that the output level of maize would decline as the use of agrochemicals increased. The negative relationship may result from the wrong application of the agrochemicals. For example excessive use of agrochemicals could lead to a decline in yield. There is also a negative relationship between seed as explanatory variable and maize yield as dependent variable. However, this relationship is not significant.

The significant and positive relationship between land and equipment as explanatory variables and maize yield as the dependent variable in the estimation results of the Rank-based Cobb-Douglas regression suggests that an increase in each of these variables will lead to an increase in the output of maize. These results are consistent with the OLS estimation. Similarly, the significant but negative relationship between the use of agrochemicals (weedicides, pesticides, fungicide and insecticide) and seed as explanatory variables and maize yield as dependent variable in the Rank-based regression suggests that the output level of maize would decline as the use of agrochemicals and seed are increased. For example excessive use of agrochemicals and seeds could lead to a decline in yield. In effect if the seeds used by farmers are higher than the recommended seed rate, yield will decline. This may lead to overcrowding which makes seedlings compete for nutrients, space and air. This result is consistent with the studies by Battese and Hassan (1999).

A comparison of the estimation result from the Cobb-Douglas model using the least squares method and the rank-based regression approach indicates that the estimates obtained in the alternative methods are similar. These results are consistent with Kloke and Mckean (2015) who demonstrated that the rank-based regression output was similar to that of the linear model and can be interpreted in the same way.

The productivity level of the farmers were examined by investigating their output elasticities and returns to scale. If the farmers increase input (labour, equipment, agrochemicals, fertilizer and seed) by one percent output changes by less than one percent whilst if farmers increase input (land) by one percent output increases by more than one percent. Noticeably, land which is positive and significantly related to output had the highest elasticity. This suggest that increasing land used in maize production will lead to increases in maize output. The importance of land in production is also noted by Rahman, Wiboonpongse, Sriboonchitta and Chaovanapoonphol (2009). On the basis of the rank based estimation, a one percentage increase in the use of agrochemicals and seed reduces output by 0.13 and 0.18 percent respectively. These reduction in output may be due to incorrect application of inputs such as seeds and agrochemicals.

Noticeably, both the OLS and the rank based estimation techniques suggest that the maize farmers were exhibiting increasing returns to scale. Thus output grows more than proportionately with any increase in input. This evidence is consistent with Wu, Devadoss and Lu (2003). This means the farmers could increase output by using more of the inputs (e.g. land, equipment and fertilizer).

CONCLUSION

This paper proposes a non-parametric rank-based estimation method to modelling the Cobb-Douglas production function as an alternative to the parametric ordinary least squares estimation approach. A comparison of the result from the Cobb-Douglas model using the least squares method and the rank-based regression approach indicates that the estimates obtained in the alternative methods are similar.

On the basis of rank-based Cobb-Douglas estimation, farm inputs such as land and equipment had a significant positive effect on maize output, whilst agrochemicals and seed had a significant negative effect on output. Furthermore, the rank-based analysis suggest that the farmers were operating at an increasing returns to scale. In summary, this paper has demonstrated that the rank-based non-parametric regression offers an alternative and a useful approach to estimating production functions.

REFERENCE

Battese, G. E., & Hassan, S. Technical Efficiency of Cotton Farmers in the Vehari District of Punjab, Pakistan. Pakistan Journal of Applied Econometrics, 1999; 15:241-253.

Chang, W. H., McKean, J. W., Naranjo, J. D., & Sheather, S. J. High-Breakdown Rank Regression. Journal of the American Statistical Association, 1999; 93(445):205-219, ISSN 0162-1459.

Cobb, C. W., & Douglas, P. H. A Theory of Production. American Economic Review 18 (Supplement): 139-165.

Czekaj, T. & Henningsen, A. Using Non-parametric Methods in Econometric Production Analysis: An Application to Polish Family Farms. Paper presented at EAAE 2011 Congress on Change and Uncertainty, Challenges for Agriculture, Food and Natural Resources in Zurich, Switzerland, 2011 August 30 to September 2.

Henningsen, A. & Kumbhakar, S. C. Semiparametric stochastic frontier analysis: An application to Polish farms during transition. Paper presented at the European Workshop on Efficiency and Productivity Analysis (EWEPA) in Pisa, Italy, 2009 June 24.

Jaeckel, L. A. Estimating Regression Coefficients by Minimizing the Dispersion of the Residuals. The Annals of Mathematical Statistics, 1972; 43:1449-1458.

Jureckova, J. Non-parametric Estimate of Regression Coefficients. The Annals of Mathematical Statistics, 1971; 42:1328-1338.

Kloke, J. & McKean, J. W. Nonparametric Statistical Methods Using R. New York: CRC Press, 2015. ISBN-13:978-1-4398-7343-4.

McKean, J. W., & Hettmansperger, T. A Robust Analysis of the General Linear Model Based on one Step r-estimates. Biometrika, 1978; 65(3):571.

Rahman, S., Wiboonpongse, A., Sriboonchitta, S., & Chaovanapoonphol, Y. Production Efficiency of Jasmine Rice Producers in Northern and North-Eastern Thailand. Journal of Agricultural Economics, 2009; 60: 419-435.

Wu, S., Devadoss, S., & Lu, Y. Estimation and Decomposition of Technical Efficiency for Sugarbeet Farms. Applied Econometrics, 2003; 35: 471-484.