European Journal of Science, Innovation and Technology
ISSN: 2786-4936
www.ejsit-journal.com
Volume 4 | Number 2 | 2024

Parametric versus Non-Parametric Models in Stochastic Frontier Analysis: A Theoretical Review

Williams Awuma[1], Isaac Appiah Amankwa[2], Akentara Richard[3], Isaac Danso[4]
[1,2]Department of Economics,
Faculty of Economics and Business Administration,
Catholic University of Ghana, Fiapre-Sunyani, Ghana
[3]Department of Accounting and Finance,
Faculty of Economics and Business Administration,
Catholic University of Ghana, Fiapre-Sunyani, Ghana
[4]Department of Accounting,
Faculty of Business Administration,
Valley View University, Kumasi, Ghana

ABSTRACT

The stochastic frontier analysis has focused on parametric models in which, the parametric functional form of the frontier function is specified, and the method involved estimating the parametric of the frontier function as well as the technical inefficiency component. This approach imposed specific distributional assumptions on the error component and applied the method of Maximum Likelihood Estimation (MLE). The content of this paper is mainly a theoretical review of the literature to extend the frontier of knowledge, particularly on parametric and non-parametric models. In this paper, we first examined the contents of the parametric model followed by the distributional-free approach which does not depend on any specific assumptions. Following the parametric half-normal model, it is observed that, the inefficiency effect, which is treated as a non-negative truncated zero-mean normal distribution and folded zero-mean normal distribution provided similar results and can be used in applied production economics with no theoretical problem. In what follows with the free approach, the Corrected Ordinary Least Squares (OLS), which used residuals provided straightforward inefficiency effect estimates from a one-sided error term compared with the traditional parametric model, and hence its flexibility did not require any underlying distribution of the dataset to be defined in advance. Given the OLS estimates, the non-parametric model can adapt to the distribution of data (data-driven) of the 196 daily farmers, making it particularly useful, when there is little or no prior information about the distribution of data which may not fit well with the parametric models. Given the sample moment-based statistic for the skewness test (skewness test on OLS residuals), if the estimated result follows the expected sign, then the rejection of the null hypothesis provided evidence for the existence of the one-sided error. To help shift the stochastic frontier analysis to more robustness owing to outliers and non-normal error distribution, researchers are encouraged to adopt the non-parametric model, as it is not constrained by a specific functional form of the error term, but determined by the data itself, thus allowing for a more data-driven modelling approach, compared with the parametric model which may not offer the best fit to the data.

Keywords: Stochastic frontier, Half-normal model, Maximum Likelihood, Skewness

INTRODUCTION

Production functions have often been applied to examine what the proportion of any increase in output over time can be, given increases in factors of production, the existence of increasing returns to scale, and what is known as 'technical progress' (Arrow et al., 1961).

While nothing has been said about what determines the proportions in which factor inputs are engaged to generate a given output, the stochastic frontier analysis pioneered by Aigner et al. (1977) and Meeusen and van de Broeck (1977) marked the departure from the traditional production functions to estimate technical efficiency of decision-making units. When modelling production behaviour, the standard production theory implicitly assumes that all production activities are on the frontier of feasible production set, subject to random noise. The common denominator is that, there is a potential maximum or minimum optimum level that defines the frontier. This is so because, the function may be saddled with disturbances on one hand showing technical efficiency and on the other, the marginal productivity representing economic efficiency (Kumbhakar & Bokusheva 2009). A production plan is therefore technically efficient, if a higher level of output is technically attainable for the given inputs (output-oriented measure) or the observed output can be produced using fewer inputs (input-oriented measure) (Kumbhakar, 1988). In the context of the stochastic frontier analysis, we understand that any lower output within the market can be attributed to the presence of inefficiencies, which are of particular interest to economists, so that policy can be considered to improve the operation of inefficient firms particularly in competitive environment (Varian, 2009). The best way to appreciate why this analysis is important is that, globally efficiency improvement is often considered as one of the most important goals behind many economic and investment reforms. For example, liberalizing our markets to competition, removal of trade barriers, and privatization of state enterprises are all motivated by the potential for efficiency improvements at least costs. Within businesses therefore, the measure of efficiency and the effects of efficiency determinants are key in the context of production economics. Following production economics, Aigner et al. (1977) and Meeusen and van den Broeck (1977) originally estimated the frontier model in which, the error components were specified, parametric distribution imposed, the log-likelihood derived, and numerical maximization procedures applied to calculate the maximum likelihood estimates. However, recent developments in the literature proposed the application of distribution-free approaches (thick frontier approach) on the error components to relax the distribution-related rigidity (Berger & Humphrey, 1991). Clearly, the choice of distribution for the random variable which represents the inefficiency effect is the issue at stake, and understanding how the assumption-free theory is adopted would help provide some relevant information in reducing difficulties in undertaking the parametric model. The extensive application of the parametric model makes it important to enquire to what extent and in what ways it is feasible to apply distribution-free assumptions given, the frontier models. The question seems so far to have been considered only from the viewpoint of approximations around a set of assumptions. Regarding which approach is the appropriate metric and what is the evidence, we seek to understand the process by dissecting parametric and non-parametric models in stochastic frontier models for similarity analysis and descriptive statistics (statistical inference) (Schmidt & Lin, 1984; Coelli, 1995; Kumbhakar et al., 2015).

LITERATURE

The Stochastic Production Frontier Model

In the context of which firm may be achieving the maximum output (Kumbhakar et al., 2010), the stochastic production frontier model for output-oriented technically inefficiency can be constructed as follows

$$logy_i = logy_i^* - u_i$$
$$loglogy_i^* = g(x_i;\beta) + v_i$$

The component u_i represents log difference between the potential and the observed output ($u_i = logy_i^* - logy_i$),where y_i denotes a scalar of observed yield, x_i is a $(J \times 1)$ vector of observed production inputs, β is a $(J \times 1)$ vector of associated coefficients, v_i is a zero-mean error term such that $u_i \geq 0$ and it represents production technical inefficiency. If input variable x is the frontier that produces the maximum output, then it is stochastic of the zero-mean random term v_i. If at any given period, $u_i \geq 0$, what this means is that the actual or observed output is lower than the frontier output y_i^*. This relationship can be expressed.

$$loglogy_i^* = g(x_i;\beta) + \epsilon_i$$

$$\epsilon_i = v_i - u_i$$

The estimated value of u_i is the output-oriented efficiency and this reduces to

$$exp(-u_i) = \frac{y_i}{y_i^*}$$

Hence, $exp(-u_i)$ is the ratio of observed output to the maximum possible output and this is the technical efficiency of the firm bounded between 0 and 1. If we apply Farrell's concept of technical efficiency, then

$$TE_i = \frac{y_i}{g(x_i;\beta)} \epsilon (0,1)$$

where $g(x_i;\beta)$ represents the deterministic production function to estimate the underlying technology. In allowing a parametric form for output, given the error components, we compute $y_i = g(x_i;\beta) \times e^{v_i - u_i}$ and when expressed in log, we obtain $log(y_i) = log\left(g(x_i;\beta)\right) + v_i - u_i$, where v_i is the normal error distributed as $v_i \sim N(0;\sigma_v^2)$ and u_i is positive denoting inefficiency distributed as $u_i \sim N(0;\sigma_u^2)$.

PARAMETRIC MODEL

The Half-Normal Model

In statistics, a parametric model or finite dimensional model is a statistical model which belongs to a family of probability distributions that possess finite number of parameters, embodies a set of statistical assumptions concerning the generation of sample data. It assumes some finite set of parameters ω, given the parameters, the future prediction of x are independent of the observed dataset $\mathcal{D}$ such that $p(x|\omega,\mathcal{D}) = p(x|\omega)$. In this section, we explore the widely applied assumptions of the *half-normal model*. Following from the stochastic frontier, we estimate the parametric distributional assumptions on v_i and u_i. The production stochastic frontier model with a normal distribution on v_i and a half-normal distribution on u_i are expressed, treating v_i and u_i as independent, i.e. $v_i \sim N(0,\sigma_v^2)$ and $u_i \sim N^+(0,\sigma_u^2)$. Notice that, if the inefficient effects are allowed to follow the half-normal distribution, two estimators can be derived. Firstly, we treat the inefficient effect as non-negative truncated of zero-mean normal distribution, $N^+(0,\sigma_u^2)$, where σ_u^2 is the variance of the normal distribution before truncation. Consider again that, a random variable Z has a normal distribution such that $z \sim N(\mu,\sigma_z^2)$ with a corresponding probability function modeled as $g(z)$, then the density function of the random variable z is approximated by the function (Kumbhakar et al., 2015)

$$f(z) = \frac{g(z)}{1 - \Phi\left(\frac{\alpha-\mu}{\sigma_z}\right)} = \frac{\frac{1}{\sigma_z}\emptyset(z-\mu)}{1 - \Phi\left(\frac{\alpha-\mu}{\sigma_z}\right)}$$

Suppose we assume that it is truncated from above, labelled as α so that $z \geq \alpha$, then $\emptyset$ and Φ represent the probability density and probability distribution functions respectively for z (Johnson et al., 1995). Consider that the inefficiency effect u_i can be generated if $\mu = 0$ and $\alpha = 0$, then the density function in this case is

$$f(u_i) = \frac{\frac{1}{\sigma}\emptyset\left(\frac{u_i}{\sigma}\right)}{1 - \Phi(0)} = \frac{2}{\sigma}\emptyset\left(\frac{u_i}{\sigma}\right) = 2(2\pi\sigma^2)^{-\frac{1}{2}} exp\left(-\frac{u_i^2}{2\sigma^2}\right), u_i \geq 0$$

Another way of thinking about this, is that we treat the inefficiency effect u_i as a folded zero-mean normal distribution (as the absolute value of a normal distribution). What this means is that, if a variable W is normally distributed, that is $w \sim N(\mu, \sigma_w^2)$ and Z possesses a folded normal distribution, described as $Z = |W|$ (Kumbhakar et al., 2015) then the density function is expressed as

$$f(z) = \left(\frac{1}{\sigma_w}\right)\left[\emptyset\left(\frac{z-\mu}{\sigma_w}\right) + \emptyset\left(\frac{z+\mu}{\sigma_w}\right)\right], z \geq 0$$

Following from the half-normal model, assuming that $\mu = 0$, we may summarize the above equation by saying that, the folded normal density function is the same for both formulations, that is to say, the two results are the same across the assumptions and hence either can be used in applied econometrics.

INEFFICIENCY EFFECTS AND CHOICE OF ESTIMATION

The Maximum Likelihood Estimation

Maximum Likelihood Estimation (MLE) provides a technique for generating parameters of probability distribution that best describe observed dataset, this is done by maximizing a likelihood function so that, the observed observations are most probable. This method enables either formulation of a regression path or determination of covariance matrix of estimates. If our observed data is assumed to be independent and identically (iid) sample: $X_1, X_2, \dots, X_n$, it implies that, they should either have the same probability mass function (data is discrete) or the same probability density function (data is continuous). Using notation $f(X|\theta)$ to denote the pdf, two notations are often applied. Firstly, we impose condition on θ to show that the likelihood of different values of variable X depends on values of our parameter and secondly, we use f for both discrete and continuous distributions (Yan, 2020). In case of discrete distribution, the likelihood is synonym with joint probability of observed data and with continuous distribution, likelihood is the joint probability density of observed data. Thus, the likelihood of all the data is the product of the likelihood of each dataset. Algebraically, the likelihood in this case given the parameter θ becomes

$$L(\theta) = \prod_{i=1}^{n} f(X_i|\theta)$$

The Likelihood-Maximization Problem

Suppose that, the random variables $X_1, \dots, X_n$ form a random sample from a distribution $f(x|\theta)$: if X is continuous stochastic variable, $f(x|\theta)$ is the pdf and if X is discrete stochastic variable, $f(x|\theta)$ is the pmf. Using a given symbol to indicate that the distribution also depends on a parameter θ, the θ could be a real-valued unknown parameter or a vector of parameters. For observed random sample $x_1, \dots, x_n$, we define $f(x_1, \dots, x_n|\theta) = f(x_1|\theta), \dots, f(x_n|\theta)$. If $f(x|\theta)$ is the pdf, $f(x_1, \dots, x_n|\theta)$ is the joint density function; if $f(x|\theta)$ is pmf, $f(x_1, \dots, x_n|\theta)$ is the joint probability. Now we refer to $f(x_1, \dots, x_n|\theta)$ as the likelihood function. We note that the likelihood function depends on the unknown parameter θ and is often described by $L(\theta)$. Assuming that, the observed random sample $x_1, \dots, x_n$ comes from discrete distribution and if an estimate of θ should be selected, we will not consider any value of θ for which it would have been impossible to obtain the dataset $x_1, \dots, x_n$ which, has been observed. Further consider that, the probability $f(x_1, \dots, x_n|\theta)$ of obtaining the actual observed data $x_1, \dots, x_n$ is high if θ has a given value i.e. $\theta = \theta_0$ and is small for every other value of θ, we will estimate the value to be θ_0. Suppose that, our sample comes from a continuous distribution, it would again be natural to find a value of θ for which the probability density $f(x_1, \dots, x_n|\theta)$ is large and use this value as an estimate of θ. For any given observed data $x_1, \ldots, x_n$, we are reminded by this principle to consider a value of θ for which, the likelihood function $L(\theta)$ is maximum and then use this value as an estimate of θ. The meaning of maximum likelihood is as follows. We select the parameter which makes the likelihood of having the obtained data at hand, maximum. With discrete distributions, the likelihood is the same as the probability. We select the parameter for the density which, maximizes the probability of the data coming from it. If we have observed values, then the estimate takes a particular numerical value which becomes the maximum likelihood estimator. The MLE is about how to maximize the likelihood function $L(\theta)$ with respect to the unknown parameter θ. Maximizing $L(\theta)$is equivalent to maximizing log $L(\theta)$) because log is a monotonic increasing function. We define $logL(\theta)$) as log likelihood function, denoted as $l(\theta)$. By maximizing $l(\theta)$ with respect to θ yields the ML estimation of the form using the approximation (Zheng, 2018)

$$l(\theta) = logL(\theta) = log\prod_{i=1}^{n} f(X_1|\theta) = \sum_{i=1}^{n} logf(X_1|\theta)$$

ESTIMATION METHODS

Normal Half-Normal Model by MLE

In the following section, we provided detailed discussion on the ML estimation with distributional assumptions on u_i and proceed as follows. Maintaining for independence of the error terms v_i (normal error distributed, we express the joint density as a product of individual functions. In such a case

$$f(u,v) = f(u) \times f(v) = \frac{2}{\sigma_u \sigma_v 2\pi} exp\left(-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right)$$

To be able to estimate the density of the error components: $\epsilon = v - u$, we would calculate the joint density of $f(u, \epsilon)$. We denote this by

$$f(\epsilon) = \int_0^\infty f(u,\epsilon)du = \frac{2}{\sigma}\phi(\epsilon\sigma^{-1})\,[1 - \Phi(\epsilon \curlywedge \sigma^{-1})]$$

$$\curlywedge = \sigma_u/\sigma_v \text{ and } \sigma^2 = \sigma_u^2 + \sigma_v^2.$$

In obtaining the density distribution of ϵ, we may rewrite the above equation as

$$E(\epsilon) = E(v-u) = E(-u) = -\frac{\sqrt{2}}{\sqrt{\pi}}\sigma_u$$

Therefore, a good estimator for the variance (ϵ) of the total observation becomes.

$$Var(\epsilon) = \sigma_\epsilon^2 = Var(u) + Var(v) = \left(\frac{\pi - 2}{\pi}\right)\sigma_u^2 + \sigma_v^2$$

The estimator for the variance of the total $[\sigma^2 = \sigma_u^2 + \sigma_v^2]$ value, relies on data coming from simple random sampling (the observations are i.i.d.) Thus, the essential assumption of the robust variance estimator is that, the observations are independently selected from the same population (Huber, 1967). Finally, the log-likelihood function based on $v_i \sim N(0; \sigma_v^2)$ and $u_i \sim N^+(0; \sigma_u^2)$ becomes

$$InL(\epsilon|\curlywedge, \sigma^2) = nIn\left(\frac{1}{\sigma}\right) + \sum_{i=1}^{n} In[1 - g\Phi(\epsilon_i \curlywedge \sigma^{-1})] - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\epsilon_i^2$$

$$\epsilon_i = logy_i - \beta logx_i$$

Having computed the estimates, the variance components are recovered for conducting hypothesis tests. In general

$$\hat{\beta}, \hat{\sigma}^2 = \hat{\sigma}_u^2 + \hat{\sigma}_v^2 \text{ and } \hat{\curlywedge} = \hat{\sigma}_u/\hat{\sigma}_v$$

$$\hat{\sigma}_v^2 = \frac{\hat{\sigma}^2}{1+\hat{\lambda}^2},\ \hat{\sigma}_u^2 = \hat{\sigma}^2 - \frac{\hat{\sigma}^2}{1+\hat{\lambda}^2}$$

Estimates of Individual Inefficiencies

In the following section, we first introduce the inefficiency effects, this is followed by detailed discussion on ML estimation. Two random variables, $v_i \sim N(0;\sigma_v^2)$ and $u_i \sim N^+(0;\sigma_u^2)$ are identified by imposing parametric distribution (Aigner et al., 1977; Meeusen & van den Broeck, 1977). Once the distributional assumptions are made, the log-likelihood function of the model is derived and numerical maximization procedures are used to obtain the ML of the model parameters. The technically inefficiency ratio TE_i can be obtained as exponential conditional expectation of $-u$, given the error component ϵ: $TE_i = e^{E(u_i|\epsilon_i)}$. On the basis of the conditional density of u given ϵ, the model which we are estimating is expressed as

$$f(u|\epsilon) = \frac{f(u,\epsilon)}{f(\epsilon)} = \frac{1}{\sigma^*\sqrt{2\pi}} = exp\left(-\frac{(u-u^*)^2}{2\sigma^{*2}}\right)\left[1-\Phi\left(-\frac{u^*}{\sigma^*}\right)\right]^{-1}$$

The distribution of u conditional on ϵ is expressed as $N^+(\mu^*,\sigma^*)$ so that

$$\mu^* = -\frac{\epsilon\sigma_u^2}{\sigma^2} = -\epsilon\gamma, \sigma^{*2} = \frac{\sigma_u^2\sigma_v^2}{\sigma^2} = \sigma^2\frac{\sigma_u^2(\sigma^2-\sigma_u^2)}{(\sigma^2)^2} = \sigma^2\gamma(1-\gamma)$$

Where $\gamma = \sigma_u^2/(\sigma_u^2+\sigma_v^2)$ is the fraction of the variance of the inefficiency to the total variance and given the distribution of $u|\epsilon$, the expected value, $E(u|\epsilon)$ may be applied as the point estimator for u_i (Jondrow et al., 1982). Thus,

$$\hat{u}_i = E(u|\epsilon) = \left(\frac{\sigma\lambda}{1+\lambda^2}\right)\left(z_i + \frac{\phi(z_i)}{\Phi(z_i)}\right), z_i = \frac{-\epsilon\lambda}{\sigma}$$

Where ϕ and Φ are defined respectively for the standard normal. In order to obtain firm-specific technical inefficiencies from $exp[-E(u|\epsilon)]$, Battese and Coelli (1988) proposed alternative estimator. In this case

$$\widehat{TE}_i = E(\exp(-u_i)\,|\epsilon_i) = \left[\Phi\left(\frac{u_i^*}{\sigma_*} - \sigma_*\right)/\Phi\left(\frac{u_i^*}{\sigma_*}\right)\right] exp\left(\frac{\sigma_*^2}{2} - u_i *\right)$$

$$u_i^* = -(logy_i - x,\beta)\sigma_u^2/\sigma^2 \text{ and } \sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$$

We note that

$$exp[-E(u|\epsilon)] \neq E\big(exp(-u_i|\epsilon_i)\big)$$

Although both estimators are unbiased, they are inconsistent because $Var(\widehat{u_\iota}) \neq 0$ for $N \to \infty$

Non-Parametric Model

Non-parametric models involve techniques which do not rely on data belonging to any particular family of probability distribution but involve methods which are *distribution-free*, which do not rely on assumptions that, the data are drawn from a given parametric family of probability distributions and the statistics is defined to be a function of a sample without] dependency on a parameter (Corder et al., 2014; Hollander et al., 2014)

Corrected Ordinary Least Squares

From the aforementioned approach, Kumbhakar et al. (2015) specified the following deterministic frontier model.

$$Iny_i = Iny_i^* - u_i, u_i \geq 0$$
$$Iny_i^* = f(x_i;\beta)$$

We note that, this model excluded the random error v_i, because it represents shocks outside the control of the firm and it is not likely to be related to the inefficiency effects u_i and therefore non-stochastic (Forsund & Hjalmarsson, 1978). Separating the intercept from the rest of the function, the following model was adopted

$$Iny_i = \beta_0 + \tilde{x}_i^{/}\tilde{\beta} - u_i$$

Where, the x_i is a vector of factor inputs including uncontrolled climate variables, which, may be in logs (log-linear function) either formulated in Cobb-Douglas or cross-product terms as with translog form. The idea is to obtain estimated frontier function bound observations (Iny_i) from above, by generating consistent estimates of the slope coefficients and the estimated production function shifted upward, such that the function after the adjustment bounds all the observations below. The model is estimated using the following procedures: Firstly, the ordinary least squares is conducted on Iny_i, using a constant of one and the following model obtained

$$Iny_i = \hat{\beta}_0 + \tilde{x}_i^{/}\hat{\beta} + \hat{e}_i$$

where $\hat{e}_i$ represent the corresponding residuals and we obtain the zero-mean OLS regression residuals $\hat{e}_i$ as

$$\hat{e}_i = Iny_i - \left[\hat{\beta}_0 + \tilde{x}_i^{/}\hat{\beta}\right]$$
$$\hat{e}_i > 0, \hat{e}_i = 0\ \hat{e}_i < 0$$

In what follows, the OLS intercept is then adjusted upwards by amount of $max\{\hat{e}_i\}$ in order that the adjusted function bounds observations from above. Following from this (Kumbhakar et al., 2015) the residuals is modelled as

$$\hat{e}_i = Iny_i - \left\{\left[\hat{\beta}_0 + max\{\hat{e}_i\}\right] + \tilde{x}_i^{/}\hat{\beta}\right\} \leq 0$$

where $\left\{\left[\hat{\beta}_0 + max\{\hat{e}_i\}\right] + \tilde{x}_i^{/}\hat{\beta}\right\}$ is the estimated function and the inefficiency effects are estimated by the equation

$$\hat{u}_i \equiv -(\hat{e}_i - max\{\hat{e}_i\}) \geq 0$$

such that $\hat{u}_i$ becomes the estimated inefficiency for $Iny_i = \beta_0 + \tilde{x}_i^{/}\tilde{\beta} - u_i$. Therefore, technical efficiency of each observation can then be computed as

$$TE_i = exp(-\hat{u}_i) \text{ for } Iny_i = \beta_0 + \tilde{x}_i^{/}\tilde{\beta} - u_i$$

Skewness Test on Ordinary Least Squares Residuals

Schmidt and Lin (1984) favoured an OLS residual test to determine the validity of the model's stochastic frontier specification, to serve as pre-test before maximum likelihood estimation is conducted. The principle of the test is that, given a production stochastic frontier model, $Iny_i = f(x_i, \beta) + \epsilon_i$ and $\epsilon_i = v_i - u_{i,}, u_{i,} \geq 0$,where ϵ_i is the composed error term and v_i distributed symmetrically around mean zero, the residuals from the corresponding OLS estimation must skew to the left, reflecting that the slope coefficients of the OLS estimation are consistent with those of the corresponding stochastic frontier model, regardless of the distribution for $u_{i,}$ after the pre-testing. Thus, a test of the null hypothesis of no skewness as opposed to the alternative hypothesis can then be conducted using the OLS residuals. Given that, the estimated skewness possesses the expected sign, the rejection of the hull hypothesis provides a rule of thumb for the presence of one-sided error. In the aforementioned scenario, Schmidt and Lin (1984) also proposed a sample-moment related statistic for skewness test, the $\sqrt{\beta_i}$ test statistic. Thus,

$$\sqrt{\beta_i} = \frac{m_3}{m_2\sqrt{m_2}}$$

where m_2 and m_3 represent the second and the third sample moments of OLS residual respectively. Thus

$$m_2 = \frac{\sum(x-\bar{x})^2}{n}, m_3 = \frac{\sum(x-\bar{x})^3}{n}$$

In this case $\sqrt{\widehat{\beta_\iota}} < 0$ would then provide evidence that the OLS residuals are within range. The distribution of $\sqrt{\beta_i}$ is assumed to be nonstandard and its critical values are tabulated in studies including D'Agostino & Pearson (1973). In addition to the direction of skewness, Coelli (1995) however suggested a variant of this test and observed that, under the null hypothesis of no skewness, m_3 of the OLS residuals is asymptotically distributed as a normal random variable with mean 0 and variance $6m_2^3/N$. It can be shown that

$$M3T = \frac{m_3}{\sqrt{6m_2^3/N}}$$

which possess an asymptotic distribution of a standard normal random variable. The advantage of the above test statistic is that, the critical values of the distribution are available and can be recovered. It should be noted that, in the context of the $\sqrt{\beta_i}$ test the one-sided error specification which represents technical inefficiency can be estimated for each observation using the ML estimation without imposing a specific distributional assumption. The distribution-free assumption otherwise referred as the *thick frontier approach* (Berger & Humphrey, 1991), which requires a full specification of the model's parameters, has improved the stochastic frontier analysis with consistent results, since the residual test serves as a screening device in this case.

EMPIRICAL LITERATURE

Thus, in this section we examine how empirically the two approaches are estimated. The paper relied on the dataset of Kumbhakar et al. (2015), which contained data on 196 dairy farms including the amount of milk produced (output) and the inputs including labour hours, feed, number of cows and land size of the farms for the illustrations of this paper.

Non-Parametric Model

Assessing the (Likely) Presence of Skewness

The table below shows evidence of a negative skewness using the point estimate statistic $\sqrt{\beta_1}$ to obtain the summary statistic of the OLS residuals e. The statistic labelled "skewness" in column 4 of Table1 below, with a value equal -0.7377269. The negative sign shows that, the OLS residuals are extremely skewed to the left as consistent with a production frontier model.

Table 1: OLS Residuals

	Residuals			
	Percentiles	Smallest		
1%	-0.4861149	-0.5450444		
5%	-0.2731016	-0.4861149		
10%	-0.1894513	-0.3645601	Obs	196
25%	-0.0821035	-0.3447073	Sum of Wgt.	196
50%	0.0101231		Mean	-3.86e-10
		Largest	Std Dev	0.1439801
75%	0.0994077	0.2340004		
90%	0.1711118	0.2554969	Variance	0.0207303
95%	0.2116255	0.2558538	Skewness	-0.7377269
99%	0.2558538	0.2701932	Kurtosis	3.92576

Source: Kumbhakar et al., 2015

In order to establish otherwise statistical significance of this statistic, Kumbhakar et al. (2015) applied the unaltered test proposed by D'Agostino et al. (1990) – a test for normality based on skewness and another based on kurtosis and then combines the two tests into an overall test statistic. The skewness/kurtosis tests for normality are presented in Table 2 below.

Table 2: Skewness/Kurtosis Tests for Normality

				…………… joint	……………..
Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	Chi2(2)	Prob> chi2
e	196	0.0001	0.0258	20.62	0.0000

Source: Kumbhakar et al., 2015

A positive value for the kurtosis (peakedness) with a value equal to 0.0258 is within limits as proposed by Bryne (2010), that is, obtaining normality thresholds between -7 to +7 is acceptable and this indicates a distribution more peaked than normal. From the normality test, since the p-value is less than 0.01, we conclude that the null hypothesis of no skewness is rejected, that is there is evidence of left-skewed error distribution which is statistically different from zero. The p-value provides evidence that, there is no need to consider the specification of the model and we proceed to estimate the stochastic frontier model. The M3T statistic proposed by Coelli (1995) was computed to confirm the rejection of the null hypothesis of no skewness on the OLS residuals. On the basis of the estimated statistic which is equal to -4.216, the null hypothesis is rejected.

Table 3: Non-Parametric Model: Standard OLS Estimation of the Model

lny_i	Coefficient	Std Err	t	ρ	[95% Conf	Interval]
$lnlabour$	0.1254299	0.0501422	2.50	0.013	0.0265262	0.224336
$lnfeed$	0.1677741	0.0433321	3.87	0.000	0.0823031	0.253245
$lncattle$	0.7710345	0.0664727	11.60	0.000	0.6399196	0.9021493
$lnland$	0.0193328	0.0448032	0.43	0.667	-0.0690398	0.1077055
$Constant$	7.272442	0.5551692	13.10	0.000	6.177392	8.367492

Source: Kumbhakar et al., 2015

Parametric Model

Half-Normal Model

In the second estimation of the stochastic frontier model, a half-normal distribution for one-sided error term was assumed and following results were obtained as shown in Table 4 below.

Table 4: Parametric Model: Half-Normal Model

lny_i	Coefficient	Std Err	t	ρ	[95% Conf	Interval]
$lnlabour$	0.102653	0.0427101	2.40	0.016	0.0189427	0.1863632
$lnfeed$	0.155628	0.0372683	4.18	0.000	0.0825836	0.2286725
$lncattle$	0.7546799	0.0574825	13.13	0.000	0.6420163	0.8673435
$lnland$	0.0360424	0.0386583	0.93	0.351	-0.0397265	0.1118114
$Constant$	7.725265	0.478893	16.13	0.000	6.786651	8.663878
σ_u	-3.133122	0.2187722	-14.32	0.000	-3.561908	-2.704336
σ_v	-5.336009	0.445821	-12.06	0.000	-6.204353	-4.468564
σ_u^2	0.435815	0.0095344	4.75	0.000	0.0283846	0.0669147
σ_v^2	0.004815	0.0021311	2.26	0.024	0.0020224	0.0114638

Source: Kumbhakar et al., 2015

The estimated coefficients (input variables) on the frontier function in Table 4 appear to be very close to those of the OLS estimates in Table 3. All inputs variables are positively related to output and statistically significant at the 5 percent except land. The output elasticity of cattle in the OLS estimation is about 77% higher than a value equal to 75% in the parametric model. Note that, in these models the returns to scale are greater than 1 indicating increasing returns and the estimates for the OLS model is larger compared with the parametric model. The estimation results provided beneath Table 4 show that the variance parameters were parameterized as exponential function equation. Hence the estimated values of the variance parameters of σ_v^2 and σ_u^2 are obtained by $\widehat{\sigma}_v^2 = \exp(-5.336009) = 0.0048$ and $\widehat{\sigma}_u^2 = \exp(-3.133122) = 0.044$.

A Likelihood Ratio (LR) Test of Inefficiency

Given a half-normal model, the LR test is used to test the null hypothesis $\sigma_u^2 = 0$, from the stochastic frontier model, the one-sided error specification which absorbed the technical inefficiency. The existence of this sided-error was tested and if no evidence of the one-sided-error is observed then the OLS estimation is adopted. Kumbhakar et al. (2015) used the LR test statistic $-2[L(H_0) - L(H_1)]$ with a 1 degree of freedom because $\sigma_u^2 = 0$ and obtained the following results.

Table 5: Critical values of the mixed chi-square distribution significance level

Df	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.455	1.642	2,705	3.841	5.412	6.635	9.500

Source: Table 1, Kodde and Palm (1986, Econometrica)

The table shows that the critical value of the statistic at the 1 percent significance level equal to 5.412 and given the model's test statistic of 16.426, the null hypothesis on technical inefficiency was rejected (Kumbhakar et al., 2015).

CONCLUSION

The difference between the approaches is that the parametric includes inefficiency and makes distributional assumptions on the error components. And if there is no inefficiency, the stochastic frontier model reduces to be estimated using the least squares method with free-assumption (normality assumption) on the error vector. This result then enables us to conduct the likelihood ratio test in order to investigate the presence of inefficiency, so that the null hypothesis of no inefficiency may be rejected. The non parametric on the other hand includes inefficiency and makes no assumptions. It provides the formular for constructing the residual bound for the inefficiency index, skewness and kurtosis tests for normality upon which the rejection of the null hypothesis provides evidence for the existence of the one-sided error. Following from this, there is no need to consider the specification of the model. Thanks for the larger flexibility of the non-parametric model, which provided straightforward estimates of the one-sided error compared with parametric model in ML. Thus, one might proceed by projecting the alternative for providing robust estimates and the processes of pre-test statistic for skewness of OLS residuals before ML estimation. From the non-parametric model, it thus appears the parametric model is not optimum (lack of fit) although in practice, it may provide quite good coverage for their distribution of observed values. The findings of this paper draw attention to the vital importance of adopting the non-parametric model owing to its flexibility, in particular obtaining larger input estimates, compared with the parametric model as shown in Tables 3 and 4 respectively. There is a new understanding that, the free-approach can speed up the estimation processes by providing more robust results. The consistent results may be as a result of the fact that parameter estimates obtained by the least squares method possess certain optimal properties and the computational procedures are fairly simple compared with other econometric techniques. Seizing these comparative analyses, the paper also draw attention to test statistic to test the skewness of the OLS residuals and providing a "rule of thumb" for rejecting the null hypothesis. The paper concluded that, a more flexible free approach reduces the computational challenges associated with the parametric and the proposed tests statistics with available critical values also guarantee strong acceptance and application of the alternative approach owing to drawing quick reference with respect to inefficiency effects.

REFERENCES

Aigner, D., Lovell, C.A. K., & Schmidt, P. (1977). Formulation and Estimation of Stochastic Frontier Production Function Models. *Journal of Econometric*, *6*, 21-37.

Allen, R.G.D. (1938). *Mathematical Analysis for Economists*. Macmillan.

Arrow, K. J., Chenery, H.B., Minhas, B.S & Solow, R.M (1961). Capital-Labour Substitution Economic Efficiency. *Review of Economics and Statistics*, *43*, 225-50.

Azzalini, A. (1985). A Class of Distribution Which Includes the Normal Ones. *Scandinavian Journal of Statistics*, *12*, 171-8.

Behr, A., & Tente, S. (2008). *Stochastic Frontier Analysis by Means of Maximum Likelihood and the Method of Moments*. Deutsche Bundesbank, Eurosystem.

Berger, A.N., & Humphery, D. B. (1991). The Dominance of Inefficiency over Scale and Product Mix Economies in Banking. *Journal of Monetary Economics*, *28*, 117-48.

Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge/Taylor & Francis Group.

Coelli, T. (1995). Estimation and Hypothesis Test for a Stochastic Frontier Function: A Monte Carlo Analysis. *Journal of Productivity Analysis*, *6*, 247-68.

Corder, G. W. & Foreman, D. I. (2014). *Nonparametric Statistics: A Step-by-Step Approach.* Wiley.

D'Agostino, R.B., & Pearson, E.S. (1973). Test for Departure from Normality. Empirical Results for the Distribution of B2 and $\sqrt{\beta_1}$. *Biometrika*, *60*, 613-22.

D'Agostino, R. B., Belanger, A. J., & D'Agostino, Jr., R. B. (1990). A suggestion for using powerful and informative tests of normality. *American Statistician*, *44*, 316–321. https://doi.org/10.2307/2684359.

Forsund, F.R., & Hjalmarsson, L. (1987). *Analyses of Industrial Structure: A Putty Clay Approach*. Stockholm: Almqvist & Wilsell.

Forsund, F.R., Hjalmarsson, L., & Summa, T. (1996). The Interplay Between Micro-Frontier and Sectorial Short-Run Production Functions. *The Scandinavian Journal of Economics*, *98*, 365-86.

Gebraeel, N.Z., Lawley, M.A., Li, R., & Ryan, J.K. (2005). Residual-Life Distribution from Components Degradation Signal: A Bayesian Approach. *IIE Transactions*, *37*(6), 543-557. https://doi.org/10.1080/07408170590929018.

Hollander, M., Wolfe, D.A., & Chicken, E. (2014). *Nonparametric Statistical Methods*. John Wiley & Sons.

Huber, P, J (1967). The Bahaviour of Maximum Likelihood Estimates under Non-standard Conditions. In *Proceedings of the Fifth Berkeley Symposium in Mathematical Statistics and Probability*, *(1)*, 221-233.

Jondrow, J., Lovel, C.A.K., Materov, I.S., & Schmidt, P. (1982). On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Models. *Journal of Econometrics*, *19*, 233-8.

Kodde, D.A. & Palm, F. C. (1986). Wald Criteria for Jointly Testing Equality and Inequality Restrictions. *Econometrica*, *54*, 1243-1248.

Kumbhakar, S. C. (1987). The Specification of Technical and Allocative Inefficiency Using Stochastic Frontier Function. *Journal of Econometrics, 34*, 335-48.

Kumbhakar, S.C. (1988). On the Estimation of Technical and Allocative Inefficiency Using Stochastic Frontier Functions: The Case of U.S. Class 1 Railroads. *International Journal Review, 29*, 727-43.

Kumbhakar, S.C., & Bokusheva, R. (2009). Modelling Farm Production Decision under an Expenditure Constraint. *European Review of Agricultural Economics, 36*, 343-67.

Kumbhakar, S. C., Wang, H. J., & Horncastle, A. (2010). Estimation of Technical Inefficiency in Production Frontier Models Using Cross-Sectional Data. *Indian Economic Review, 45*, 1-77.

Kumbhakar, S.C., Wang, H.J., & Horncastle, A. (2015). *Practitioner's Guide to Stochastic Frontier Analysis Using Stata*. Cambridge University Press.

Meeusen, W., & van den Broeck, J. (1977). Efficiency Estimation from Cobb-Douglas Production Function with Composed Error. *International Economic Review, 18*, 435-44.

Plancade, S., Rozenholc, Y., & Lund, E. (2012). Generalization of the Normal-Exponential Models: Exploration of a more Accurate Parametrization for the Signal Distribution on Illumina BeadArrays. *Bioinformation, 13*, 329.

Schmidt, P., & Lin, T.F. (1984). Simple Tests of Alternative Specification in Stochastic Frontier Model. *Journal of Econometrics*, *24*, 349-61.

Sheskin, D. J. (2003). *Handbook of Parametric and Nonparametric Statistical Procedures*. CRC Press.

Varian, H.R. (2009). *Microeconomic Analysis*. W.W. Norton & Company.