CHAPTER 3

LITERATURE REVIEW ON EFFICIENCY, MEASUREMENT AND EMPIRICAL APPLICATIONS

3.1 Introduction

The objective of this chapter is to give an overview of the concept of efficiency and frontier models, the different approaches to its measurement in the context of frontier models and empirical studies on efficiency. Approaches to efficiency measurement are broadly specified into parametric and non parametric approaches. Given the large volume of theoretical and empirical literature in the field of efficiency measurement, the review of empirical studies is further subdivided into three namely: a review of empirical comparative studies in agriculture, a review of empirical comparative studies in other sectors where the distance function approach was used and finally a review of empirical studies in Nigerian agriculture. The review is intended not only to provide a proper understanding of the specific area of research but it also helps the researcher to establish a vivid framework to be employed for analysis.

3.2 The Concept of Efficiency and Frontier Models

In microeconomic theory a production function is defined in terms of the maximum output that can be produced from a specified set of inputs, given the existing technology available to the firms involved (Battese, 1992). The maximum possible output becomes relevant in order to answer certain economic questions such as the measurement of efficiency of firms, hence the introduction of frontier production functions which estimates the maximum output as function of inputs. Similarly, a cost frontier function would give the minimum cost as a function of output quantity and input prices.

The papers by Debreu (1951) and Koopmans (1951) mark the origin of discussion on the measurement of productivity and efficiency in the economic literature. The work of Debreu and Koopmans was first extended by Farrell (1957) in order to perform the measurement of productivity and efficiency. The productivity of an economic agent can be measured simply as a scalar ratio of outputs to inputs that the agent uses in its production process. Productivity could be measured either as partial productivity such as yield per hectare (land productivity) or output per person (labour productivity) or more appropriately as total factor productivity (TFP) which is defined as ratio of aggregate outputs to aggregate inputs. An economic agent's productivity may vary based on differences in production technology, in the efficiency of the production process, in the environment in which production occurs, and finally in the quality of inputs used by the agent (Haghiri, 2003). On the other hand, efficiency is measured by comparing observed and optimal values of the agent's outputs and inputs. Prior to Farrell's work, efforts were made to measure efficiency by interpreting the average productivity of inputs, then to construction of efficiency indexes. However, these methods were found unsatisfactory by economists and agricultural economists as the methods suffered from one shortcoming to another. The use of the traditional least squares methods for estimating the production function has been critiqued as this is not consistent with the definition of the production function. The estimated functions could at best be described as average or response functions because such regression estimates the mean output (rather than the maximal output) given quantities of inputs (Schmidt, 1986). This led to the development of a better-founded theoretical method for measuring efficiency, i.e. the frontier method. Frontiers models are described as bounding functions (Coelli, 1995b).

The frontier approach holds a number of advantages over average or response functions as well as over non-frontier models. There are two main benefits that result from estimating frontier functions, as compared to estimating average functions using ordinary least squares (OLS) approach. First, when a frontier function is estimated, the result is strongly influenced by the best performing firm, and therefore the frontier reflects the technology set that the most efficient firm employs. However, the estimation of an average function only reflects the technology set employed by an average firm. Second, frontier functions provide a useful performance benchmark. These functions normally represent best practice technology, against which the efficiency of other firms within the industry can be measured. Frontier models also provide a number of advantages over non-frontier models like the one proposed by

Lau and Yotopoulos (1971). A non-frontier model yields efficiency measures for groups of firms, whereas a frontier model can provide firm specific efficiency measures to the researcher. Another advantage of the frontier methodology is that the word 'frontier' is consistent with the theoretical definition of a production, cost, and profit function, i.e., a solution to a maximum and minimum problem. These advantages make the frontier methodology popular in applied economic research (Forsund et al., 1980; Bravo-Ureta and Pinherio, 1993; Haghiri, 2003; Alene, 2003).

Frontier functions can be classified based on certain criteria. First, based on the way the frontier is specified, frontiers may be specified as parametric function of inputs or non-parametric. Second, it may be specified as an explicit statistical model of the relationship between observed output and the frontier or it may not. Finally, a frontier function can be classified according to how one interprets the deviation of a group of agents or firms from the best performing agents in the sample. In this sense, frontier functions can be either deterministic or stochastic. In the sub-sections that follow, we broadly classify the frontier models into parametric or non-parametric frontiers.

3.3 Non-Parametric Frontier Approach

A non-parametric approach neither specifies a functional form for the production technology nor makes an assumption about the distribution of the error terms. In other words it is robust with respect to the particular functional form and to the distribution assumptions. The non-parametric approach is mainly deterministic in nature. In a deterministic production frontier model, output is assumed to be bounded from above by a deterministic (non-stochastic) frontier. However, the possible influence of measurement errors and other statistical noise upon the shape and positioning of the estimated frontier is not accounted for.

The original work of Farrell (1957) serves an important starting point for discussion of non-parametric frontiers. Farrell illustrated the measurement of efficiency using an input-oriented approach. His argument is embodied in figure 3.1. This illustration was done by considering a firm using two inputs x_1 and x_2 to produce output y, such that the production frontier is $y = f(x_1, x_2)$ Assuming constant returns to scale , then one can write $1 = f(x_1 / y, x_2 / y)$, that is the frontier technology can be characterized by a unit isoquant and this is denoted SS' in figure 3.1. Knowledge of the unit isoquant of a fully efficient firm permits the measurement of technical efficiency. For a given firm using (x_1*, x_2*) defined by point A $(x_1*/y, x_2*/y)$ to produce a unit of output $y*$, the ratio OQ/OA measures technical efficiency and it defines the ability of a firm to maximize output from a given set of inputs. The ratio measures the proportion of (x_1, x_2) needed to produce $y*$. Technical efficiency takes a value between zero and one and therefore provides an indication of technical inefficiency. Thus, the technical inefficiency of the firm, 1-OQ/OA, measures the proportion by which (x_1*, x_2*) could be reduced (holding the input ratio x_1 / x_2 constant) without reducing output. A firm that is fully technically efficient would lie on the efficient isoquant (example, point Q) and it takes a value of 1.

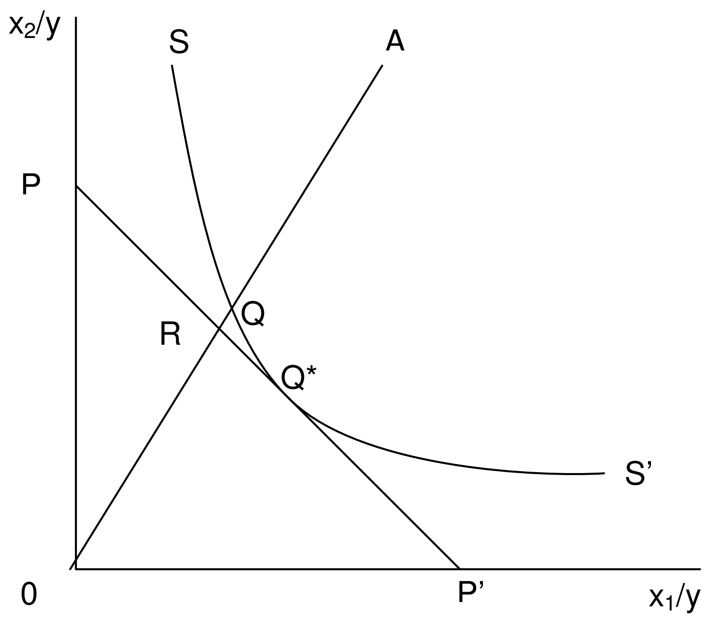

Figure 3.1: Technical, Allocative and Economic Efficiency

Further, Farrell demonstrated that the unit isoquant can provide a set of standards for measuring allocative (referred to as price efficiency by Farrell) efficiency. Let PP' represent the ratio of input prices. Then the ratio OR/OQ measures the allocative efficiency (the ability of a firm to use inputs in optimal proportions, given the respective prices at point A). Correspondingly, allocative inefficiency is 1- OR/OQ. The distance RQ is the reduction in production costs which would have been achieved had production occurred at Q*- the allocatively and technically efficient point, rather than Q- the technically efficient, but allocatively inefficient point. Finally, the ratio OR/OA measures the economic efficiency (referred to as overall efficiency by Farrell) and correspondingly 1-OR/OA measures the total inefficiency. The distance RA is the cost reduction achievable which is obtained from moving from A (the observed point) to Q* (the cost minimizing point).

In this approach, the efficient unit isoquant is not observable; it must be estimated from a sample of observations. The approach is non-parametric because Farrell simply constructs the free disposal convex hull of the observed input-output ratios by linear programming techniques which are supported by a sub-set of the sample, with the rest of the sample points lying above it.

According to Forsund et al. (1980), the major advantage of non-parametric approach is that no functional form is imposed on the data. One disadvantage of the approach is that the frontier is computed from a supporting subset of observations, and is therefore particularly susceptible to extreme observations and measurement error. A second disadvantage is that the estimated functions have no statistical properties upon which inferences can be made; however, recent developments are attempting to overcome this drawback.

Farrell's approach has been extended by Charnes et al. (1978) giving rise to what is known as data envelopment analysis (DEA). The technique envelopes observed production possibilities to obtain an empirical frontier and measures efficiency as the distance to the frontier. Efficient firms are those that produce a certain amount of or more outputs while spending a given amount of inputs, or use the same amount of or less inputs to produce a given amount of outputs, as compared with other firms in the test group. This approach generalizes Farrell's approach of computing the efficiency frontier as a piecewise-linear convex hull in the input coefficient space to multiple outputs. Charnes et al. (1978) reformulated Farrell's approach into calculating the individual input saving efficiency measures by solving a linear programming problem for each unit under the constant returns to scale (CRS) assumption while Banker et al. (1984) extended it to the case of variable returns to scale (VRS) since imperfect competition, financial constraints may cause a firm not to be operating on an optimal scale, the assumption upon which CRS is appropriate. Charnes et al. (1978) proposed a model which had an input-orientation. The DEA can be considered as a non-parametric approach to estimation of distance functions (Färe et al., 1985; 1994).

Assuming there is data on K inputs and M outputs on each of N firms. For the *i*th firm, these are represented by the vectors x_i and y_i, respectively. The K x N input matrix, X and the M x N output matrix, Y, represent the data of all N firms. The purpose of the approach is to construct a non-parametric envelopment frontier over the data points such that all observed points lie on or below the production frontier.

The input-oriented constant returns to scale DEA frontier is defined by the solution to N linear programs of the form:

$$
\begin{aligned}
&\min_{\theta,\lambda} \theta, \\
&\text{subject to } \; -y_i + Y\lambda \geq 0, \\
&\qquad\qquad\quad \theta x_i - X\lambda \geq 0, \qquad\qquad (3.1) \\
&\qquad\qquad\quad \lambda \geq 0
\end{aligned}
$$

where θ is a scalar and λ is an Nx1 vector of constants. The value of θ is an index of technical efficiency for the *i*th firm and will satisfy $0 \leq \theta \leq 1$, with value of 1 indicating a point on the frontier and hence a technically efficient firm, according to Farrell (1957) definition. Thus, $1-\theta$ measures how much a firm's inputs can be proportionally reduced without any loss in output.

However, the assumption of CRS is correct only as long as firms are operating at an optimal scale (Coelli et al, 2002). Using the CRS DEA model when firms are not operating at their optimal scale will cause the technical efficiency measures to be influenced by scale efficiencies and thus the measure of technical efficiency will be incorrect. The CRS linear programming problem can easily be modified to account for variable returns to scale by adding the convexity constraint: $N1'\lambda = 1$ to equation (3.1) to provide an input-oriented VRS model:

$$\min_{\theta,\lambda} \theta$$

$$\text{subject to } -y_i + Y\lambda \geq 0,$$
$$\theta x_i - X\lambda \geq 0, \qquad (3.2)$$
$$N1'\lambda = 1$$
$$\lambda \geq 0$$

where N1 is an Nx1 vector of ones. This approach forms a convex hull of intersecting planes which envelope the data points more tightly than the CRS conical hull and thus provide technical efficiency scores which are greater than or equal to those obtained using the CRS model.

The output-oriented models are very similar to their input-oriented counterparts. For instance, the output-oriented VRS model is defined by solution to N linear programs of the form:

$$\max_{\phi,\lambda} \phi$$

$$\text{subject to } -\phi y_i + Y\lambda \geq 0,$$
$$x_i - X\lambda \geq 0, \qquad (3.3)$$
$$N1'\lambda = 1$$
$$\lambda \geq 0$$

where $1 \leq \phi < \infty$, and ϕ is the proportional increase in output that could be achieved by the ith firm, with input held constant. $1/\phi$ defines a technical efficiency score which varies between zero and one. The CRS output-oriented model can be defined similarly by removing the convexity constraint, $N1'\lambda = 1$ from equation (3.3).

In the input-oriented models, the method sought to identify technical inefficiency as a proportional reduction in input usage. They are input-oriented because they try to find out how to improve the input characteristics of the firm concerned so as to become efficient. The output-oriented measure sought to identify technical inefficiency as a proportional increase in output production. The input and output orientations provide the same value under CRS but are unequal under the assumption of a VRS. Thus, the input- and output-oriented models will estimate exactly the same frontier and therefore, by definition, identify the same set of firms as being efficient. It is only the efficiency measures associated with the inefficient firms that may differ between the two methods. Given that linear programming cannot suffer from such statistical problems as simultaneous equation bias, the choice of an appropriate orientation is not very crucial. Essentially, one should select an orientation according to which quantities (inputs or outputs) the managers have most control over. In many instances, the choice of orientation will have only minor influences upon the scores obtained (Coelli, 1995b, Coelli and Perelman, 1999).

With availability of price information, it is possible to consider a behavioural objective, such as cost minimization or revenue maximization so that both technical and allocative efficiency can be measured. For the case of a VRS cost minimization, one would run the input-oriented DEA model set out in equation (3.2) to obtain technical efficiency (TE). One would then run the following cost minimization DEA

$$\begin{aligned} &\min_{\lambda, x_i*} \; w_i' x_i *, \\ &\text{subject to } \; -y_i + Y\lambda \geq 0, \\ &\qquad x_i * - X\lambda \geq 0, \\ &\qquad N1'\lambda = 1 \\ &\qquad \lambda \geq 0 \end{aligned} \tag{3.4}$$

where w_i is a vector of input prices for the *i*th firm and x_i* is the cost minimizing vector of input quantities for the *i*th firm given the input prices w_i and the output levels y_i and this is calculated by the linear programming. The total cost efficiency (CE) or economic efficiency of the *i*th firm would be calculated as

$$CE = \frac{w_i' x_i *}{w_i' x_i} \tag{3.5}$$

That is, the ratio of minimum cost to observed cost. One can then use equation (3.5) to calculate the allocative efficiency residually as

$$AE = \frac{CE}{TE} \tag{3.6}$$

This procedure will include any slacks into the allocative efficiency measure. This is often justified on the grounds that slack reflects an inappropriate input mix (Ferrier and Lovell, 1990).

The aim of DEA analysis is not only to determine the efficiency rate of the units reviewed, but also to find target values for inputs and outputs for an inefficient unit. After reaching these values, the unit would arrive at the threshold of efficiency. The major disadvantage of the deterministic DEA approach is that it takes no account of possible influence of measurement error and other noise in the data and as such it has been argued that it produces biased estimates in the presence of measurement error and other statistical noise. However, it has the advantage of removing the necessity to make arbitrary assumptions about the functional form of the frontier and the distributional assumption of the error term. With DEA, multiple output technologies can be examined very easily without aggregation.

As it has been stated earlier, one of the main drawbacks of non-parametric techniques is their deterministic nature. This is what traditionally has driven specialised literature on this issue to describe them as non-statistical methods. Nevertheless, recent literature has shown that it is possible to define a statistical model allowing for the determination of statistical properties of the non-parametric frontier estimators (Murillo-Zamorano, 2004). For instance, DEA models with stochastic variations have recently received attention (Banker, 1993; Land et al., 1993; Sengupter 2000a; Simar and Wilson, 1998, 2000a, 2000b; Huang and Li, 2001; Kao and Liu, 2009; Shang et al., 2009). Simar and Wilson (1998, 2000a, 2000b) for example, methodically studied statistical properties of DEA models, and developed bootstrap algorithms which can be used to examine the statistical properties of efficiency scores generated through DEA. Therefore, one might conclude that today statistical inference based on non-parametric frontier approaches to the measurement of economic efficiency is available either by using asymptotic results or by using bootstrap. However, a couple of main issues still remain to be solved, namely the high sensitivity of non-parametric approaches to extreme values and outliers, and also the way for allowing stochastic noise to be considered in a non-parametric frontier framework (Murillo-Zamorano, 2004).

3.4 Parametric Frontier Approach

The parametric approach involves a specification of a functional form for the production technology and an assumption about the distribution of the error terms. The major advantage of the parametric approach compared to the non-parametric approach is the ability to express the frontier technology in a simple mathematical form. However, the parametric approach imposes structure on the frontier that may be unwarranted. The parametric approach often imposes a limitation on the number of observations that can be technically efficient. For example, in the case of homogeneous Cobb-Douglas form, when the linear programming algorithm is used, there will in general be only as many technically efficient observations as there are parameters to be estimated (Forsund et al, 1980). This approach can be subdivided into deterministic and stochastic frontiers. The parametric deterministic approach is further subdivided into statistical and non-statistical methods.

3.4.1 Deterministic Non-Statistical Frontiers

Few people adhered to the non-parametric approach by Farrell (1957). Almost as an after thought, Farrell (1957) proposed a second approach. In this approach, Farrell proposed computing a parametric convex hull of the observed input-output ratios. He recommended the Cobb-Douglas production function for this purpose given the limited selection of functional form then. He acknowledged the undesirability of imposing a specific (and restrictive) functional form on the frontier but also noted the advantage of being able to express the frontier in a simple mathematical form. This suggestion was however not followed up by Farrell.

Aigner and Chu (1968) were the first to follow Farrell’s suggestion. In order to express the frontier in a mathematical form, they specified a Cobb-Douglas production frontier, and required all observations to be on or beneath the frontier. Their model may be written as:

$$\ln y_i = \ln f(x_i;\alpha) - u_i; \qquad u \geq 0 \tag{3.7}$$

where y_i is the output of the *i*th sample firm; x_i is the inputs of the *i*th firm, u_i is a one-sided non-negative random variable associated with firm-specific factors that contribute to the *i*th firm inability to attain maximum efficiency of production. The one sided error term, u_i forces $y \leq f(x)$. The elements of the parameter vector, α, may be estimated either by linear programming (minimizing the sum of the absolute values of the residuals subject to the constraint that each residual is non-positive) or by quadratic programming (minimizing the sum of squared residuals, subject to the same constraint). Although Aigner and Chu (1968) did not do so, the technical efficiency of each observation can be computed directly from the vector of residuals, since u represents technical efficiency.

A major problem with this approach is that it produces estimates that lack statistical properties. That is, the programming procedure produces estimates without standard errors, t-ratios, etc. This is because no statistical assumptions are made about the regressors or the disturbance term in equation (3.7) and therefore inferences cannot be obtained.

3.4.2 Deterministic Statistical Frontiers

The previous models were critiqued on their lack of statistical properties. This problem can be addressed by making some assumptions about the disturbance term. The model in equation (3.7) can be written as

$$\ln y = f(x)e^{-u}, \tag{3.8}$$

or

$$\ln y = \ln[f(x) - u], \tag{3.9}$$

where $u \geq 0$, implying $0 \leq e^{-u} \leq 1$, $\ln[f(x)]$ is linear in the Cobb-Douglas case presented in equation (3.7). Some assumptions are usually made about u and x and that is, that u are independently and identically distributed (iid), with mean μ and finite variance and that x is exogenous and independent of u. Any number of distributions for u (or e^{-u}) could be specified. Aigner and Chu (1968) did not explicitly assume such a model though it seems clear it was assumed implicitly. However, the first to explicitly propose this type of model was Afriat (1972), who proposed a two-parameter beta distribution for e^{-u}, and that the model be estimated by maximum likelihood method. This amounts to gamma distribution for u, as considered further by Richmond (1974). On the other hand Schmidt (1976) has demonstrated that if u is exponential, then Aigner and Chu's linear programming procedure is maximum likelihood, while their quadratic programming procedure is maximum likelihood if u is half-normal.

In the frontier setting, there are some problems with maximum likelihood. First, maximum likelihood estimates (MLE) depend on the choice of distribution for u such that different assumptions yield different estimates. This is a problem because there are no good *a priori* arguments for choice of any particular distribution. Second, the range of the dependent variable (output) depends on the parameters to be estimated (Schmidt, 1976). This is because $y \leq f(x)$ and $f(x)$ involves the parameters which are to be estimated. For any one-sided error distribution, $y \leq f(x)$ violates one of the usual regularity conditions for consistent and asymptotic efficiency of maximum likelihood estimators (namely, that the range of the random variable should not depend on the parameters). Thus, the statistical properties of the MLE's are in general uncertain. Greene (1980a) finds sufficient conditions on the distribution of u for the MLE's to have their usual desirable asymptotic properties: (i) if g is the density of u, g(0) = 0, i.e. the density of u is zero at $u = 0$ and (ii) $g'(u) \rightarrow 0$ as $u \rightarrow 0$, i.e. the derivative of the density of u with respect to its parameters approaches zero as u approaches zero. However, as Schimdt (1986) noted, it is clearly not desirable that one's assumptions about the error term be governed by the need to satisfy such conditions.

An alternative method of estimation based on ordinary least squares was first proposed by Richmond (1974) and is called corrected OLS or COLS. Suppose equation (3.9) is assumed to be linear (Cobb-Douglas) and letting μ be the mean of u, then

$$\ln y = (\alpha_0 - \mu) + \sum_{i=1}^{n} \alpha_i \ln x_i - (u - \mu) \qquad (3.10)$$

where the new error term has zero mean. Since the error term satisfies all the usual ideal conditions except normality, equation (3.10) can be estimated by OLS to obtain best linear unbiased estimates of $(\alpha_0 - \mu)$ and of α_i. If a specific distribution is assumed for u, and if the parameters of the distribution can be derived from higher-order (second, third, etc.) central moments, then these parameters can be consistently estimated from the moments of the OLS residuals. Since μ is a function of these parameters, it can also be estimated consistently, and this estimate can be used to correct the OLS constant term, which is consistent estimate of $(\alpha - \mu)$. Thus, COLS provides consistent estimates of all the parameters of the frontier. However, this technique poses some difficulties. First, some of the residuals may still have wrong signs after correcting the constant term so that these observations end up above the estimated production frontier. This makes COLS seem not to be a very good technique for computing technical efficiency of individual observations. There are two ways of resolving this problem namely, by use of stochastic frontier approach or to estimate equation (3.10) by OLS, then correct the constant term not as above, but by shifting it up until no residual is positive, and one is zero. Another difficulty with COLS technique is that the correction to the constant term is not independent of the distribution assumed for u. That is, different assumptions yields systematically different corrections for the constant term, and systematically different estimates of technical efficiency, except for the special case var (u) =1. However, this problem again can be resolved by shifting the function upward until no residual is positive, and one is zero.

3.4.3 Stochastic Frontiers

They emerged as an improvement over average functions and deterministic frontiers. In the deterministic frontiers, all variations in the firm performance are attributed solely to variation in firm efficiencies relative to the common family of frontiers, be it production, cost or profit frontiers. Thus, the idea of a deterministic frontier shared by all firms ignores the very real possibility that a firm's performance may be affected by factors that are entirely outside its control such as bad weather, input supply breakdowns etc as well as factors under its control (inefficiency). To lump these effects of exogenous shocks, both fortunate and unfortunate, together with the effects of measurement error and inefficiency into a single one-sided error term, and to label the mixture inefficiency is questionable and is a major weakness of deterministic frontiers.

Forsund et al. (1980) noted that this conclusion is reinforced if one considers also the statistical noise that every empirical relationship contains. The standard interpretation is that first, there may be measurement error on the dependent variable. Second, the equation may not be completely specified with the omitted variables individually unimportant. Both of these arguments hold just as well for production functions as for any kind of equation, and it is dubious at best not to distinguish this noise from inefficiency, or to assume that noise is one-sided. It is on this basis that the stochastic frontier (composed error) model was independently proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977). The vital idea behind the stochastic frontier model is that the error term is composed of two parts. A symmetric component permits random variation of the frontier across firms, and captures the effects of measurement error, other statistical noise, and random shocks outside the control of the firm. A one-sided component captures the effects of inefficiency relative to the stochastic frontier.

The stochastic frontier function may be defined according to Battese (1992) as:

$$y_i = f(x_i, \alpha)\exp(\varepsilon_i), \;\; i = 1,.....N \qquad (3.11)$$

where $\varepsilon_i = v_i - u_i$. (3.12)

The stochastic frontier is $f(x_i, \alpha)\exp\ (v_i)$, y_i is the output of the ith firm and is bounded above by the stochastic quantity, x_i are the inputs of the ith firm. ε_i is a random variable. v_i is the random error having zero mean, and is associated with random effects of measurement errors and exogenous shocks that cause the deterministic kernel $f(x_i, \alpha)$ to vary across firms. Technical inefficiency is captured by the one-sided error component $\exp(-u_i)$, where $u_i \geq 0$ implying that all observations must lie on or beneath the stochastic production frontier.

The random errors, v_i were assumed to be independently and identically distributed as $N(0, \sigma^2{}_v)$ random variables and independent of the u_i's , which were assumed to be non-negative truncations of the half-normal distribution i.e., $\left|N(0, \sigma_u{}^2)\right|$ or exponential distribution i.e. EXP$(\mu, \sigma_u{}^2)$. Aigner et al. (1977) considered half-normal and exponential distributions but Meeusen and van den Broeck (1977) considered exponential distribution only. Stevenson (1980) has shown how the half-normal and exponential distributions can be generalized to truncated normal ($N(\mu, \sigma_u{}^2)$) and gamma distributions, respectively. There was a tendency for researchers to use the half-normal and truncated normal distributions probably because of ease of estimation and interpretation and more so, as there were no standard tests for distribution selection. However, Lee (1983) proposed a Lagrange-Multiplier test to assess different distributions for the inefficiency term. Given the assumptions of the stochastic frontier model (3.11), inference about the parameters of the model can be based on the maximum likelihood estimators because the standard regularity conditions are satisfied.

Technical efficiency of an individual firm is defined in terms of the ratio of the observed output to the corresponding frontier output, conditional on the levels of inputs used by that firm. Thus, the technical efficiency of firm i in the context of the stochastic production function expressed in equations (3.11) and (3.12) is given as

$$TE_i. = \frac{y_i}{y_i^*} = \frac{y_i}{f(x_i;\alpha)\exp(v_i)^*} = \exp(-u_i) \qquad (3.13)$$

The prediction of technical efficiencies of individual firms associated with the stochastic frontier production function (3.11) was considered impossible until the appearance of Jondrow et al. (1982). Following Jondrow et al. (1982) and Battese and

Corra (1977) reparameterization, the firm specific technical efficiency can be predicted by the conditional expectation of the non-negative random variable, u_i, given that the random variable, ε_i, is observable. The technical efficiency of the *i*th firm is then given by:

$$E(u_i / \varepsilon_i) = \frac{\sigma_u \sigma_v}{\sigma} \left[\frac{f(\cdot)}{1 - F(\cdot)} - \frac{\varepsilon_i}{\sigma} \left(\frac{\gamma}{1-\gamma} \right)^{1/2} \right] \tag{3.14}$$

where ε_i are the estimated residuals for each firm, $f(\cdot)$ and $F(\cdot)$ are the values of the standard normal density function and standard normal distribution function, respectively, evaluated at $\frac{\varepsilon_i}{\sigma}\left(\frac{\gamma}{1-\gamma}\right)^{1/2}$. The parameters of the model, i.e. α, $\sigma^2 = \sigma_v{}^2 + \sigma_u{}^2$ and $\gamma = \sigma_u{}^2 / \sigma^2$ can be obtained from the maximum likelihood estimation of equation (3.11). γ is bounded between zero and one and it explains the total variation of output from the frontier which can be attributed to technical inefficiency. The estimates of v_i and u_i can be obtained by substituting the estimates of ε_i, γ, and σ. Thus, the technical efficiency of individual firms can be measured as $TE_i = \exp(-E(u_i / \varepsilon_i)$ which represents the level of technical efficiency of the *i*th firm relative to the frontier firm. However Battese and Coelli (1988) derived the best predictor of TE given as $E(-u_i / \varepsilon_i) = \left[\frac{1 - F(\sigma_A + \gamma\varepsilon_i / \sigma_A)}{1 - F(\gamma\varepsilon_i / \sigma_A)} \exp(\gamma\varepsilon_i + \sigma_A^2 / 2) \right]$.

One can test whether any form of stochastic frontier production is needed at all by testing the significance of the γ parameter. If the null hypothesis, that γ equals zero, is accepted, this would indicate that $\sigma_u{}^2$ is zero and hence that the u_i should be removed from the model, leaving a specification with parameters that can be consistently estimated using ordinary least squares (Coelli, 1996a).

There are two approaches to estimating the inefficiency effect models, that is, the second part of the stochastic frontier models that provides explanation for variation in efficiency of firms. These may be estimated with either a one step procedure or a two step procedure. In a one step procedure estimates of all the parameters are obtained in one step. The inefficiency effects are defined as a function of the firm specific factors (as in the two-stage approach) but they are then incorporated directly into the MLE. That is, both the production frontier and the inefficiency effect models are estimated simultaneously. For the two-step procedure, the production frontier is first estimated and the technical efficiency of each firm is derived. These are subsequently regressed against a set of variables, z, which are hypothesized to influence the firms' efficiency. The two-stage procedure has been critiqued of inconsistency in the assumptions about the distribution of the inefficiencies. This is because in the first stage, the inefficiencies are assumed to be independently and identically distributed (iid) in order to estimate their values. However, in the second stage, the estimated inefficiencies are assumed to be a function of a number of firm specific factors, and hence are not identically distributed unless all the coefficients of the factors are simultaneously equal to zero (Coelli, et al. 1998, Herrero and Pascoe, 2002). Thus, the distributional assumptions used in either step contradict each other (Coelli, et al, 2005). Kumbhakar et al. (1991) argued that the estimated technical coefficients and technical efficiency indices are biased when the determinants of technical efficiency are not included in the first step of the regression. They provided a one-step procedure which determines the influence of socioeconomic variables on technical efficiency while estimating technical coefficients of the production frontier. Kalirajan (1991), on the other hand, has defended the practice of the two-step regression on the basis that socioeconomic variables have a roundabout effect on production.

Although the two-step procedure is critiqued of producing biased results, there seems to be little evidence on the severity of this bias. For example, Caudill and Ford (1993) provide evidence on the bias of the estimated technological parameters, but not on the efficiency levels or their relationship to the explanatory variables. However, Wang and Schmidt (2002) identified two sources of bias namely, that the first step of the two-step procedure is biased for the regression parameters if the z and the inputs, x_i are correlated. Secondly, that even if z and x are independent, the estimated inefficiencies are under-dispersed when the effect of z on inefficiency is ignored. This causes the second-step estimate of the effect of z on inefficiency to be biased downward (toward zero). Therefore, they suggested that a one step procedure be employed to overcome this problem. There appear to be no consensus in the literatures on the use of either one step or two step procedure and the choice may be solely that of the analyst.

The Cobb-Douglas functional form is the commonly used in estimating the stochastic production frontier. Although its most attractive feature is simplicity, but this is associated with a number of restrictions. Most notably the returns to scale are restricted to take the same value across all firms in the sample, and elasticities of substitution are assumed equal to one. However, more flexible functional forms like the translog production function have also received attention. The translog form imposes no restriction upon returns to scale or substitution possibilities, but has the drawback of being susceptible to multicollinearity and degrees of freedom problems (Coelli, 1995b). In any case, the choice of appropriate function form can be made by conducting a likelihood ratio test between competing models.

Stochastic frontier analysis (SFA) has both advantages and disadvantages. The advantages include first, it controls for random unobserved heterogeneity among the firms. The inefficiency effect can be separated from statistical noise. With non-parametric methods, any deviation of an observation from the frontier must be attributed to inefficiency, which makes the results very sensitive to outliers or measurement errors and uncertainty. Second, by using SFA, the statistical significance of the variables determining efficiency can be verified using statistical tests, though this is also true for recent bootstrapped DEA models. Third, the firm specific inefficiency is not measured in relation to the "best" firm, as it is done in non-parametric approaches. Hence, SFA is again less sensitive to outliers in the sample. Disadvantages of the SFA approach consist of the need for distributional assumptions for the two error components as well as the assumption of independence between the error terms and the regressors. Further, implementation of the model requires the choice of an explicit functional form, the appropriateness of which raises questions.

The stochastic frontier specification has been altered and extended in a number of ways. These extensions include: consideration of panel data and time-varying technical efficiencies, the extension of the methodology to cost, revenue and profit frontiers, estimation of stochastic input and output distance functions, the estimation of systems of equations, the decomposition of the cost frontier to account for both technical and allocative efficiency. A review of most of these extensions is provided by Forsund et al. (1980), Schmidt (1986), Bauer (1990), and Coelli (1995b). However, in the subsequent sub-sections brief explanations of some these extensions are given.

3.4.3.1 Panel Data

Cross sectional data provides a snapshot of producers and their efficiency. Panel data provides more reliable evidence on their performance, because they enable one to track the performance of each producer through sequence of time periods. In the Panel data model, a time varying or time invariant inefficient effect may be specified. Also, the model may assume either a fixed or random effect. A significant advantage of panels is that given consistently large time periods, they permit consistent estimation of the efficiency of individual producers, whereas the Jondrow et al. (1982) technique does not generate consistent estimators in a cross-sectional context (Kumbhakar and Lovell, 2000). Another advantage of the panel data is that the distributional assumptions about the efficiency term upon which stochastic frontier rely is no longer necessary. Also the assumption of independence between the inefficiency term and input levels is unnecessary with panel data. Again, panel data increases degrees of freedom for estimation of parameters and it permits the simultaneous estimation of technical change and technical inefficiency changes over time. However, the dearth of panel data on farmers especially in developing country agriculture has constrained the use of panel data methodologies.

3.4.3.2 Duality Considerations and Cost System Approaches

The consideration of duality extends not only to cost minimization but also profit maximization, though cost minimization is often made in the dual frontier literatures. Thus, the discussion here is basically on cost minimization behaviour. It is very simple to change the sign of the inefficiency error component u_i and convert the stochastic production frontier model to a stochastic cost frontier model such that we have:

$$C_i = c(y_i, w_i; \beta).\exp(v_i + u_i) \tag{3.15}$$

where C_i is the cost of production of the *i*th firm, $c(y_i, w_i; \beta).\exp(v_i)$ is the stochastic cost frontier, w_i is a vector of input prices of the *i*th firm, y_i is output of the *i*th firm; β is an vector of unknown parameters; v_i are random variables which are assumed to be independently and identically distributed $N(0, \sigma_v{}^2)$ and independent of , u_i, which are non-negative random variables which are assumed to account for the cost of inefficiency in production, which are often assumed to be iid $\left|N(0, \sigma_u{}^2)\right|$. In this cost function, the u_i now defines how far the firm operates above the cost frontier. If allocative efficiency is assumed, then u_i is closely related to the cost of technical efficiency. If this assumption is not made, the interpretation of the u_i in a cost frontier is less clear, with both technical and allocative inefficiencies possibly involved (Coelli, 1996a). The Jondrow et al. (1982) technique may be used to provide an estimate of the overall cost inefficiency, but the difficult remaining problem is to decompose the estimate of u_i into estimates of the separate costs of technical and allocative inefficiency. Schmidt and Lovell (1979) accomplished the decomposition for the Cobb-Douglas case while Kopp and Diewert (1982) obtained the decomposition for the more general translog case based on deterministic frontier.

According to Coelli (1995b) there are basically three reasons for considering the alternative of dual forms of the production technology, such as the cost or profit function. First, is to reflect alternative behavioural objectives such as cost minimization. Second is to account for multiple outputs. Third, is to simultaneously predict both technical and allocative efficiency. The choice of whether to estimate a production or cost frontier may be based on exogeneity assumptions. It is more natural to estimate a production frontier if inputs are exogenous and a cost frontier if output is exogenous (Schmidt, 1986). Schmidt and Lovell (1979) suggested a maximum likelihood system estimation of their Cobb-Douglas frontier, involving the cost function and k-1 factor demand equations as this is expected to improve the precision of the parameter estimates. Such a system can be specified as follows:

$$\ln y_i = A + \sum_j \alpha_j \ln x_{ij} + v_i - u_i \tag{3.16}$$

$$\ln x_{i1} - \ln x_{ij} = \ln p_{ij} - \ln p_{i1} + \ln \alpha_1 - \ln \alpha_j + \varepsilon_{ij}, \qquad j = 2, \ldots k \tag{3.17}$$

$$\ln c_i = K + \frac{1}{r}\ln y_i + \sum_{j=1}^{k} \frac{\alpha_j}{r} \ln p_{ij} - \frac{1}{r}(v_i - u_i) + (E_i - \ln r) \tag{3.18}$$

where y is output, x's inputs, p's are prices, i indexes firms and j indexes inputs. Equation (3.16) is a stochastic production frontier, while equation (3.17) is the set of first order conditions for cost minimization. Equation (3.18) is the cost function. ε_{ij} represents allocative efficiency. $r = \sum_{j=1}^{k} \alpha_j$ is the returns to scale, E_i (equation 3.19) is given as a function of ε's and the parameters. The cost of technical inefficiency is $\frac{1}{r}u_i$, while the cost of allocative inefficiency is $(E_i - Inr)$. The latter is non-negative, and zero if $\varepsilon_{ij} = 0$ for all j.

$$E_i = \sum_{j=2}^{n} \frac{\alpha_j}{r}\varepsilon_{ij} + In\left[\alpha_1 + \sum_{j=2}^{n} \alpha_j e^{-\varepsilon_{ij}}\right] \tag{3.19}$$

This approach faces two serious draw backs. First, in some cases it may not be practical or appropriate to estimate a cost frontier. For instance, it will not be practical to estimate a cost function when input prices do not vary among firms and it will not be appropriate when there is a systematic deviation from cost-minimising behaviour in an industry. Second, Schmidt and Lovell (1979) systems estimation and the technical and allocative efficiency measurement are limited to self-dual functional forms like the Cobb-Douglas. Once one specifies a more flexible functional form like the translog forms which are not self-dual, a problem arises. The major problem with employing a translog form is associated with how to model the relationship between the allocative inefficiency error which appears in the input share equations and that which appears in the cost function (sometimes referred to as the 'Green Problem' because it was first noted by Green (1980b). Although a number of approaches have been suggested and applied in modelling the Greene problem ranging from analytic solution (e.g. Kumbhakar, 1989), approximate solution (e.g. Schmidt, 1984) to qualitative solution (e.g. Greene 1980b), debate still continues on how best to address this problem. Coelli (1995b) noted that a sound approach to take (given that the cost minimizing assumption is appropriate and suitable price data are available) is to estimate the cost function using single equation maximum likelihood method and then use the method proposed by Kopp and Diewert (1982), and refined by Zeischang (1983) for deterministic frontier case or that extended by Bravo-Ureta and Rieger (1991) for stochastic frontier case following the primal route, to decompose the cost efficiencies into their technical and allocative components. If the Cobb-Douglas functional form is considered appropriate, then the procedure involved simplify to those which are outlined in Schmidt and Lovell (1979). Berger (1993) found that efficiency estimates using no cost share equations, partially restricted share equations, and fully restricted share equations gave very similar efficiency results.

3.4.3.3 Production Frontier and Efficiency Decomposition

Given that it may not be appropriate to estimate a cost function when there is little or no variation in prices among sample firms, Bravo-Ureta and Rieger (1991) developed an alternative approach to decompose the cost efficiency into technical and allocative efficiencies. They followed a primal route in their methodology. The methodology involved using the level of output of each firm adjusted for statistical noise, the observed input ratio and the parameters of the stochastic frontier production function (SFPF) to decompose economic efficiency into technical and allocative efficiency. Then the cost function is analytically derived from the parameters of the SFPF. To illustrate the approach, a stochastic frontier production function is given as:

$$Y_i = f(X_i;\beta) + \varepsilon_i \tag{3.20}$$

$$\varepsilon_i = v_i - u_i \tag{3.21}$$

where ε_i is the composed error term. The two components v_i and u_i are assumed to be independent of each other, where v_i is the two-sided, normally distributed random error and u_i is the one-sided efficiency component with a half normal distribution. Y_i is the observed output of the ith firm, X_i is the input vectors of ith firm and β is unknown parameters to be estimated. The parameters of the SFPF were estimated using the maximum likelihood method. Subtracting v_i from both sides of the equation (3.20) results in

$$Y_i^* = Y_i - v_i = f(X_i;\beta) - u_i \tag{3.22}$$

where Y_i^* is the observed output of the *i*th firm adjusted for statistical noise captured by v_i. From equation (3.22), the technically efficient input vector, X_i^T, for a given level of Y_i^* is derived by solving simultaneously equation (3.22) and the input ratios, $X_1 / X_k = \rho_k (k > 1)$, where ρ_k is the ratio of the observed inputs.

Assuming the production function is self-dual function like the Cobb-Douglas production function, the corresponding dual cost frontier can be derived and written in a general form as:

$$C_i = h(W_i, Y_i^*; \delta) \tag{3.23}$$

where C_i is the minimum cost of the *i*th firm associated with output Y_i^*; W_i is a vector of input prices of the *i*th firm; and δ is a vector of parameters which are functions of the parameters in the production function.

The economically efficient (cost minimizing) input vector, X_i^E, is derived by using Shephard's Lemma and then substituting the firm's input prices and adjusted output quantity into the system of demand equations:

$$\frac{\partial C_i}{\partial W_i} = X_i^E (W_i, Y_i^*; \delta) \tag{3.24}$$

For a given level of output, the corresponding technically efficient, economically efficient and actual costs of production are equal to $W_i X_i^T$, $W_i X_i^E$ and $W_i X_i$, respectively. These three cost measures are then used as the basis for calculating the technical and economic (cost) efficiency indices for the *i*th firm :

$$TE_i = \frac{W_i X_i^T}{W_i X_i} \tag{3.25}$$

and

$$EE_i = \frac{W_i X_i^E}{W_i X_i} \tag{3.26}$$

Following Farrel (1957), allocative efficiency can be calculated by dividing economic efficiency (EE) by technical efficiency (TE):

$$AE_i = \frac{W_i X_i^E}{W_i X_i^T} \tag{3.27}$$

3.4.3.4 Distance Functions and Efficiency Decomposition

The production, cost, profit and perhaps revenue functions are well known alternative methods of describing a production technology. These functions have been used by economists to measure efficiencies. Of recent the application of distance functions is growing. The majority of recent distance function studies have been motivated by a desire to calculate technical efficiencies or shadow prices. The principle advantage of the distance function representation is that it allows the possibility of specifying a multiple-input, multiple-output technology when price information is not available or alternatively when price information is available but cost, profit or revenue function representations are precluded because of violations of the required behavioural assumptions (Coelli and Perelman 2000). The distance function contains the same information about technology as does the cost function but may have some advantages econometrically over the cost function if, for example, input prices are the same for firms, but input quantities vary across firms (Bauer, 1990).

The output distance function measures how close a particular level of output is to the maximum attainable level of output that could be obtained from the same level of inputs if production is technically efficient. In other words, it represents how close a particular output vector is to the production frontier given a particular input vector (Mawson et al., 2003). The definition of an output-distance function starts with a definition of the production technology of the firm using the output set, $P(x)$, which represents the set of all output vectors, $y \in R_+^M$, which can be produced using the input vector, $x \in R_+^K$. That is,

$$P(x) = \{ y \in R_+^M : x \text{ can produce } y \} \tag{3.28}$$

The output-distance function is then defined on the output set, $P(x)$, as

$$D_O(x, y) = \min\{\theta : (y/\theta) \in P(x)\} \tag{3.29}$$

$D_O(x, y)$ is non-decreasing, positively linearly homogeneous and convex in y, and decreasing in x (Lovell et al., 1994). The distance function, $D_O(x, y)$, will take the value which is less than or equal to one if the output vector, y, is an element of the feasible production set, $P(x)$. That is, $D_O(x, y) \leq 1$ if $y \in P(x)$. Furthermore, the distance function will take the value of unity if y is located on the outer boundary of the production possibility set. That is,

$$\begin{aligned} D_O(x, y) = 1 \text{ if } y \in \text{ Isoq } P(x) \\ = \{y : y \in P(x), \omega y \notin P(x), \omega > 1\}; \end{aligned} \tag{3.30}$$

A Stochastic Output Distance Function (SODF) is not the same as a Stochastic Frontier Production Function (SFPF). Both consider the maximum feasible output from a given set of inputs. The difference is that SODF is defined in a set theoretic framework which involves vector of outputs and inputs and can only be implemented empirically by normalizing using one of the outputs whereas SFPF is simply defined for the case of one output or aggregated outputs and does not require normalization.

An input-distance function is defined in a similar manner as the output distance function. However, rather than looking at how the output vector may be proportionally expanded with the input vector held fixed, it considers by how much the input vector may be proportionally contracted with the output vector held fixed. The input-distance function may be defined on the input set, $L(y)$, as

$$D_I(x, y) = \max\{\rho : (x / p) \in L(y)\} \tag{3.31}$$

where the input set, $L(y)$, represents the set of all input vectors, $x \in R_+^K$, which can produce the output vector, $y \in R_+^M$. That is,

$$L(y) = \{\ x \in R_+^K :\ x \text{ can produce y}\} \tag{3.32}$$

$D_I(x, y)$ is non-decreasing, positively linearly homogenous and concave in x, and increasing in y. The distance function, $D_I(x, y)$, will take a value which is greater than or equal to one if the input vector, x, is an element of the feasible input set, $L(y)$. That is, $D_I(x, y) \geq 1$ if $x \in L(y)$. Furthermore, the distance function will take a value of unity if x is located on the inner boundary of the input set.

Under the assumption of constant returns to scale (CRS), the input distance function is equivalent to the inverse of the output distance function (i.e., $D_O = 1/D_I$) (Färe et al. 1993, 1994). That is, the proportion by which one is able to radially expand output (with input held fixed), will be exactly equal to the proportion by which one is able to radially reduce input usage (with output held constant). However, under variable returns to scale (VRS) this condition need not hold.

The distance function can be illustrated graphically. For instance, the input distance function as exemplified in Coelli et al. (2003) is shown in figure 3.2. Here two inputs, x_1 and x_2, are used to produce output y. The isoquant SS', is the inner boundary of the input set, reflecting the minimum input combinations that may be used to produce a given output vector. In this case, the value of the distance function for a firm producing output y, using the input vector defined by point P, is equal to the ratio, OP/OQ.

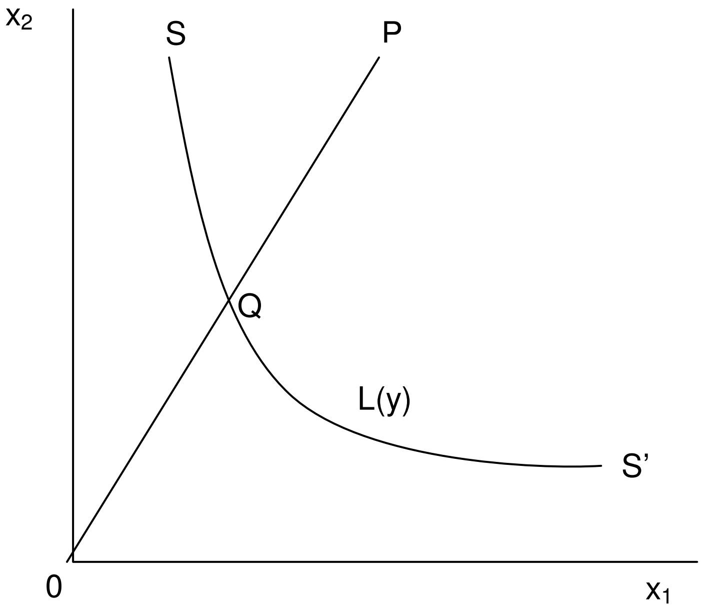

Figure 3.2: The input distance function and the input set

In empirical literatures on efficiency measurement involving distance functions, different methods have been employed to estimate the function. These include the construction of parametric frontier using linear programming methods (Färe et al., 1994; Coelli and Perelman, 1999; Alene and Manfred, 2005); the construction of non-parametric piece-wise linear frontier using the linear programming method known as data envelopment analysis (DEA) (e.g. Färe et al., 1989; Färe et al., 1994; Coelli and Perelman, 1999; Alene and Manfred, 2005); estimation of parametric frontier using corrected ordinary least square (COLS) (e.g. Lovell et al., 1994; Grosskopf et al., 1997; Coelli and Perelman, 1999) and maximum likelihood estimation (MLE) of a parametric stochastic distance frontier (e.g. Coelli et al., 2003; Irz and Thirtle, 2004; Solis et al., 2009). However all of these studies have basically focused on analysing technical efficiency except that of Coelli et al. (2003) that applied the cost decomposition approach to estimate both technical and allocative efficiency.

Decomposition of cost efficiency in a single equation stochastic input distance frontiers framework was first developed by Coelli et al. (2003) to overcome the problems that arise when one either tries to estimate a cost frontier and then use duality to derive the implicit production frontier as in Schmidt and Lovell (1979) or alternatively estimating a primal production technology, and then derive the implicit cost frontier as in Bravo-Ureta and Rieger (1991). The input distance function approach avoids all of these problems because first it does not require price information to vary among firms. Second, it is robust to systematic deviations from cost minimising behaviour. Third, it does not suffer from simultaneous equations bias when firms are cost minimisers or shadow cost minimisers (Coelli, 2000; Coelli et al., 2003). Finally, the approach has an added advantage over production function in that it can easily accommodate multiple outputs without aggregation as in production function.

A general framework of the decomposition approach is described below using the parametric input distance function. It is noted that the value of the distance function is not observed so that imposition of a functional form for $D_I(x,y)$ does not permit its direct estimation. A convenient way of handling this problem was suggested by Lovell et al. (1994) who exploit the property of linear homogeneity of the input distance function. Given a general form of a parametric input distance function as:

$$D_I = f(x,y) \tag{3.33}$$

where f is a known functional form such as Cobb-Douglas or translog. Linear homogeneity implies:

$$\lambda D_I = \lambda f(x,y) \qquad \forall \lambda > 0 \tag{3.34}$$

Assuming x is a vector of K inputs and setting $\lambda = 1/x_1$, where x_1 is its (arbitrarily chosen) first component, then equation (3.34) can be expressed in a logarithmic form as:

$$\ln(D_I / x_1) = \ln f(x/x_1, y) \tag{3.35}$$

or

$$\ln(D_I) - \ln(x_1) = \ln f(x/x_1, y) \tag{3.36}$$

and hence

$$-\ln(x_1) = \ln f(x/x_1, y) - \ln(D_I) \tag{3.37}$$

where $-\ln(D_I)$ is defined as $\varepsilon = v - u$ to indicate that the distance term may be interpreted as a traditional stochastic frontier analysis disturbance term. That is, the distances in a distance function (which are radial distances between the data points and the frontier) could be due to either noise (v) or technical inefficiency (u) which is the standard SFA error structure (Coelli et al., 2003). Therefore equation (3.37) can be rewritten as:

$$-\ln(x_1) = \ln f(x/x_1, y) + v - u \tag{3.38}$$

The random errors, v are assumed to be independently and identically distributed as $N(0, \sigma^2{}_v)$ random variables and independent of the u's , which are assumed to either be a half-normal distribution i.e., $\left|N(0, \sigma_u{}^2)\right|$ or exponential distribution i.e. EXP$(\mu, \sigma_u{}^2)$ or truncated normal (($N(\mu, \sigma_u{}^2)$) or gamma distributions. The predicted radial input-oriented measure of TE for an *i*th firm is given as:

$$T\hat{E}_i = 1/\hat{D}_I = E[\exp(u_i)|v_i - u_i] \tag{3.39}$$

In other words, $1 - T\hat{E}_i$ measures the proportion by which costs would be reduced by improving technical efficiency, without reducing output. A value greater than one for the input distance function ($\hat{D}_I$) indicates that the observed input-output vector is technically inefficient. When the producer is operating on the technically efficient frontier or the isoquant, the parametric input distance function attains a value of one. The technically efficient input quantities can be predicted as follows:

$$\hat{x}_{ji}^T = x_{ji} \times T\hat{E}_i; \qquad j = 1,2...K \tag{3.40}$$

Using the first order condition for cost minimisation, the duality between the cost and input distance function can be derived and expressed in a general form as:

$$C_i(w_i, y_i) = \underset{x}{Min}\{w_i x_i : D_I(x_i, y_i) \geq 1\} \tag{3.41}$$

where C is the cost of production and w denotes a vector of input prices. From this minimisation problem, it is possible to relate the derivatives of the input distance function to the cost function and by making use of Shephard's Lemma, allocative efficiency and cost efficiency can then be computed.

The current study makes a comparison of the production function and distance function frontier results and proposes an integrated efficiency model for resolution of model selection problems in efficiency studies and agricultural policy analysis.

3.5 Empirical Studies on Efficiency Measurement

A number of empirical studies both in agricultural and non-agricultural sectors have applied the frontier models since the pioneering work of Farrell (1957). However, given the large volume of theoretical and empirical literature in the field of efficiency measurement, a general review of comparative studies in agriculture and other sectors is provided. The review of comparative studies in other sectors is limited to those involving the use of distance functions since the application of distance functions is not vast yet. Finally, to place this section in the Nigerian context, a review of some of the efficiency studies in Nigeria that used either one or more of the frontier approaches is provided.

3.5.1 Empirical Comparative Studies in Agriculture

Ferrier and Lovell (1990) compared two techniques for estimating production economies and efficiencies with each having both advantages and disadvantages. One approach involved the econometric estimation of a cost frontier; the second was a series of linear programs which calculate a production frontier. Their results showed that the two different techniques yielded very similar results regarding cost economies, and dissimilar results regarding cost efficiencies.

Kalaitzandonakes and Dunn (1995) in their study on the relationship between technical efficiency and education calculated technical efficiency with three alternative frontier methods for a sample of Guatemalan corn farms namely, a deterministic statistical frontier (COLS), a stochastic production frontier estimated by MLE technique and a non-parametric DEA. The three alternative frontier methods resulted in significant differences both in the average technical efficiency of the sample and the efficiency rankings of individual farms. Furthermore, following two-step procedures where technical efficiency is regressed against a set of explanatory variables, it was shown that the choice of efficiency measurement technique can alter the importance of education as a contributing factor to increased technical efficiency. The study therefore recommended that inferences based on efficiency studies should be cautious as technical efficiency may not be dependable when difficulties in the empirical measurement of conceptual variables and other measurements errors are not explicitly accounted for. Hence, an alternative approach was therefore presented for investigating the relationship between education and efficiency while accounting for difficulties in the measurement of conceptual variables and measurement errors.

Sharma et al. (1997) compared the performance of stochastic and DEA production frontiers in predicting technical efficiencies for a sample of Hawaii swine producers. Under the stochastic method, the efficiency measures were estimated under the specifications of the Cobb-Douglas production frontier for which the inefficiency effects have the truncated-normal distribution. In the DEA analyses, the output-oriented frontiers were estimated under the specifications of constant and variable returns to scale. The estimated mean technical efficiency in the stochastic frontier is larger than those obtained from the DEA analyses. The correlation between the technical efficiency rankings of the two approaches was positive and highly significant.

Sharma et al. (1999) analyzed technical, allocative and economic efficiency measures derived for a sample of swine producers in Hawaii using the parametric stochastic efficiency decomposition technique and nonparametric data envelopment analysis. The results from both approaches revealed considerable inefficiencies in swine production in Hawaii. The estimated mean technical and economic efficiencies obtained from the parametric technique were higher than those from DEA for CRS models but quite similar for VRS models, while allocative efficiencies were generally higher in DEA. However, the efficiency rankings of the sample producers based on the two approaches were highly correlated, with the highest correlation being achieved for the technical efficiency rankings under CRS. Based on mean comparison and rank correlation analyses, the return to scale assumption was found to be crucial in assessing the similarities or differences in inefficiency measures obtained from the two approaches. Analysis of the role of various firm-specific factors on productive efficiency shows that farm size had strong positive effects on efficiency levels. Similarly, farms producing market hogs were more efficient than those producing feeder pigs.

Mbaga et al. (2000) measured the technical efficiency of two groups of dairy farms in Quebec. While the actual production technology was unknown, they checked three commonly used functional form (Cobb-Douglas, translog, and generalized Leontief) along with three alternative potential inefficiency distributions (half-normal, truncated-normal, and exponential). To gain information about the robustness of the obtained technical efficiency, they also estimated a production frontier using data envelopment analysis (DEA) as an alternative methodology. The authors obtained cross-sectional data on 1143 farms that specialized in dairy production in 1996. They divided these farms into two groups (non-maize and maize regions) as proxies for differences in climate and soil quality. Their results indicated that all the correlation coefficients, as well as the rank correlation coefficients between the DEA scores and those of the parametric models, were relatively low. The average efficiency scores obtained from the DEA approach were 0.9215 for the non-maize region and 0.95 for the maize region. For the maize region, the average DEA score was similar to those generated by the generalized Leontief (GL) function, but scores were somewhat lower for the non-maize region. The DEA model showed that about 66 percent of the farms were classified as being over 90 percent efficient, while more than 93 percent of the farms fell in this category with the GL function, irrespective of the efficiency distribution.

Wadud and White (2000) compared DEA and stochastic frontiers production function (SFPF) measures of the efficiency of 150 rice farmers in two villages in Bangladesh. For the stochastic frontier model both the one-stage and two-stage procedures were implemented. The technical efficiency estimates SFPF from was lower than that from CRS DEA but greater than that of the VRS DEA. Efficiency rankings were however positive and significant. Results from both approaches indicate that technical efficiency is significantly influenced by factors measuring environmental degradation and irrigation infrastructure.

Wadud (2003) assessed estimates of technical, allocative and economic efficiency of farms using farm- level survey data for rice farmers in Bangladesh. Results from the stochastic production efficiency decomposition technique and Data Envelopment Analysis were compared. Inefficiency effects were modelled as a function of farm specific human capital variables, irrigation infrastructure and environmental factors. The results from both approaches showed that there was substantial technical, allocative and economic inefficiency in production and that analysis of technical, allocative and economic inefficiency in terms of land fragmentation, irrigation infrastructure and environmental factor were robust.

Premanchandra (2002) evaluated the extent to which alternative methods of estimation vary from one another in measuring technical efficiency. Using data from the New Zealand dairy industry for the year 1993, the paper calculated farm-specific technical efficiency estimates and mean technical efficiency estimates for each estimation method. The methods adopted include the stochastic frontier production function, corrected ordinary least squares regression (COLS) and Data Envelopment Analysis. The results derived show that the mean technical efficiency of an industry is sensitive to the choice of the production frontier method. In general, the SFPF and DEA frontiers resulted in higher mean technical efficiency estimates compared to the COLS production frontier. The resulting mean TE estimate from the SFPF production frontier was significantly higher than that of DEA, except under the variable returns to scale DEA model. The results from the DEA and SFP frontiers also indicate that New Zealand dairy farmers were operating nearer to or at the efficiency frontier. All three methods are consistent in ranking individual production units in terms of technical efficiency.

Haghiri (2003) used a stochastic nonparametric frontier regression analysis to estimate and compare the technical efficiency of a large set of dairy producers in

Canada with their counterparts in the U.S. Using a panel data set, an iterative procedure called a smoothing process was used to estimate the mean response function and its parameters constructed in a generalized additive model (GAM). Using the method of locally scoring smoothing, the parameters of the regression function were estimated by employing two separate nonparametric techniques: locally weighted scatter plot smoothing (LOWESS), and spline smoothing. After estimating the response function and its parameters, the technical efficiency scores were computed. These efficiency indices were also compared with the one obtained from conducting a stochastic parametric (translog) frontier function using both the maximum likelihood estimation (assuming a half-normal distribution) and the COLS methods. The results show that the overall mean technical efficiency obtained from translog function for all regions is higher than that of the corresponding values obtained from the nonparametric approaches. Both parametric and nonparametric methodologies indicated evidence of differences between the mean technical efficiency of dairy farms in all regions meaning that various policies implemented in the two countries significantly impacted the performance of dairy producers.

Jaforullah and Premanchandra (2004) estimated technical efficiency for the New Zealand dairy industry using three different estimation techniques under both constant returns to scale and variable returns to scale in production. The approaches used were the econometric stochastic production frontier (SPF), corrected ordinary least squares (COLS) and data envelopment analysis (DEA). Mean technical efficiency of the industry was found to be sensitive to the choice of estimation technique. In general, the SPF and DEA frontiers resulted in higher mean technical efficiency estimates than the COLS production frontier.

Alene and Manfred (2005) compared the performances of the parametric deterministic distance functions (PDF) and DEA with applications to adopters of improved cereal technology in Eastern Ethiopia. Although they found positive and significant correlations between the two approaches, the result from PDF was more robust when analysis was subjected to sensitivity to possible outliers. The results from the preferred PDF approach revealed that adopters of improved technology have average technical efficiencies of 79 percent, implying that they could potentially raise their food crop production by an average 21 percent through full exploitation of the potentials of improved varieties and mineral fertilizer.

Herrero (2005) compared four different approaches data envelopment analysis, stochastic production frontier, panel data, and distance function to estimation of technical efficiency of the Spanish Trawl Fishery that was operated in Moroccan water. Their findings show that the efficiency estimates were similar and highly correlated. Thus, they conclude that none of the methodologies can be said to be better than the rest; rather, the most appropriate methodology depends on the characteristics of the production process, the degree of stochasticity, number of outputs and possibility of aggregation.

Johansson (2005) estimated technical, allocative and economic efficiency scores for an unbalanced panel of Swedish dairy farms, using data envelopment analysis and the stochastic production frontier approach. The mean technical, allocative and economic efficiency indices for the entire period were 0.55, 0.75, and 0.41, respectively in the SFPF model. However, when the data envelopment analysis was applied, the technical, allocative and economic efficiency indices were 0.74, 0.61, and 0.45, respectively. Thus, the mean technical and economic efficiency indices were higher under DEA than under SFPF whereas the reverse was the case for allocative efficiency. A paired t-test results showed that the measures of technical and economic efficiency were significantly higher under the DEA approach while allocative efficiency was higher under SFPF approach. However, both SFPF and DEA provided similar rankings. Further results showed a positive relationship between size and efficiency. Finally it was concluded that the main challenge facing the Swedish dairy farms is to enhance their cost minimizing skills.

Tingley et al. (2005) calculated technical efficiency for segments of the English Channel fisheries using the econometric stochastic production frontier (SPF) and the non-stochastic, linear-programming data envelopment analysis (DEA) methodologies. The influence of factors most affecting technical efficiency was analysed using an SPF inefficiency model and tobit regression of DEA-derived scores. While the overall DEA technical efficiency scores were affected by random error and thus lower that those of SPF, the results demonstrated that both techniques were able to produce reasonable models of factors that affect efficiency. With only one exception, the analysis of the efficiency scores using the two methods (DEA and SPF) was consistent, at least in terms of direction of the effect. They concluded that based on the explanatory power of the models and the number, sign and consistency of significant variables between models, the tobit regression of DEA-derived scores are generally as robust as those of the comparative SPF inefficiency model and therefore, tobit regression of DEA-derived technical efficiency scores can be used as an alternative method to explain inefficiency where SPF model specification is problematic.

Alene et al. (2006) analysed efficiency of intercropping annual and perennial crops in Southern Ethiopia by comparing technical efficiency predictions from parametric stochastic frontier production function (SFPF), parametric deterministic distance functions (PDF) and non-parametric DEA using different orientations. The mean technical efficiency from SFPF (72 percent) were lower than that obtained from PDF (89-93 percent) and DEA (92-94 percent). Further, SFPF gave higher technical efficiency variation across farms but efficiency rankings were similar for the three approaches. They concluded that whether stochastic or deterministic frontiers yield higher or lower estimates cannot be determined a priori. Testing the stability of technical efficiency estimates from the three approaches, they found that PDF and DEA are more robust than SFPF. Based on similarity of results from DEA and PDF, the final efficiency scores were obtained from their geometric mean.

Bojnec and Latruffe (2007) investigated the determinants of technical efficiency of Slovenian farms by comparing results from parametric stochastic frontier production function and the non-parametric data envelopment analysis. They obtained consistent results for all the included variables except for land where the two methods produced contradicting result both in terms of sign and significance. They thus concluded that the influence of land is undetermined.

Odeck (2007) compared data envelopment analysis and stochastic frontier analysis to assess efficiency and productivity growth of Norwegian grain producers. He found consistency between the approaches to the extent that there were potentials for efficiency improvements, but the magnitudes depend on the model applied and by segmentation of the data set. However, he warned that policy-makers should not be indifferent with respect to the approach used for efficiency and productivity measurement at least with respect to the magnitudes of potential for efficiency improvements and productivity growth since each approach may give different results.

3.5.2 Empirical Comparative Studies in other Sectors involving Distance Functions

Coelli and Perelman (1999) investigated technical efficiency in European railways. They compared the results obtained from three alternative methods of estimating multi-output distance functions. Specifically they considered the construction of a parametric frontier using linear programming (PLP); data envelopment analysis (DEA) and corrected ordinary least squares (COLS). Input-orientated, output-orientated and constant returns to scale (CRS) distance functions were estimated and results from these were compared. Their results indicated a strong degree of correlation between the input- and output-orientated results for each of the three methods. Significant correlations were also observed between the results obtained using the alternative estimation methods. The strongest correlations were observed between the parametric linear programming and the COLS methods. Based on similarity of results, they used the geometrical mean of efficiency scores from all model results for final ranking.

Coelli and Perelman (2000) compared results from three specifications of distance functions estimated by COLS and two specifications of single output production frontiers. The study focused on the use of technical efficiency as a measure of performance of the European railways. The results obtained indicate substantial differences in parameter estimates and technical efficiency rankings, casting significant doubt upon the reliability of the single-output models. Therefore, their final preferred model was the (unrestricted) input distance function with a mean technical efficiency level of 0.863 and mean values for individual companies that range from 0.784 for Italy to 0.980 for the Netherlands.

Jamasb and Pollitt (2003) compared 63 regional electricity distribution utilities in the six European countries. To calculate technical efficiency and to consider the effects of choosing the variables and methods, they used six DEA, two COLS, and two SFA techniques of estimating input distance functions. Their results show a strong correlation between the non-parametric base model DEA-CRS and the parametric COLS and SFA models. However, they found that the mean and minimum efficiency scores in DEA-CRS base model were significantly lower than the other two models. They also found that the DEA-CRS base model efficiency scores were significantly lower than those of corresponding DEA-VRS and that the VRS model exhibited a somewhat weaker correlation with the latter model than with COLS and SFA models.

Estache et al. (2004) applied DEA and econometric methods for performance assessment and ranking of South American electricity units. Specifically they estimated two parametric distance models (an input distance function and an input requirement function) and four deterministic nonparametric DEA models (two input distance functions, one with variable returns to scale and another with constant returns to scale, and two input requirement functions, one with variable returns to scale and another with constant returns to scale). Testing the internal consistency of results obtained from all approaches, first they found that efficiency levels from different approaches were significantly different. Secondly, they found high correlation between different econometrics as well as DEA models. However, there was low correlation between DEA and econometrics models. Thirdly, they found that the best and worst performers were identified reasonably well by all the DEA models but the selection of a particular SFA model was not a trial choice. They also tested the external consistency of different approaches by determining the year-to year stability of DEA and SFA efficiency estimates over time. The results suggest that the efficiency scores were stable over time.

Cuesta et al. (2009) compared the performance of parametric stochastic hyperbolic distance functions with DEA in the analysis of environmental efficiency of U.S. electricity generating units and found that although the means and distributions of the models were significantly different, the ranking of the units by each model is similar.

3.5.3 Recent Empirical Efficiency Studies in Nigerian Agriculture

Ajibefun (2002) analysed the determinants of technical efficiency of small scale farmers in Nigeria and the effect of policy changes on technical efficiency, using a Cobb-Douglas stochastic frontier production function. The result showed a wide variation in the estimated technical efficiencies, ranging between 0.18 and 0.91, and a mean value of 0.63, indicating a wide room for improvement in the technical efficiency. The results of simulation of policy variables showed that the level of technical efficiency significantly increased with rising level of education and farming experience.

Ogunyinka and Ajibefun (2004) analysed the determinants of technical inefficiency among the farmers that are participating in the Ondo State chapter of the National Directorate of Employment program in Nigeria. They obtained an average efficiency score of 61 percent which translates to average inefficiency of 39 percent. Employing a second stage tobit regression analysis, it was found that extension visits, higher education, land input and membership of farm association were significant factors influencing technical efficiency with only extension visit having a negative influence, while others had the expected positive influence. The study concluded that sound education, efficient inputs supply strategy and public awareness of efficient technology are key factors necessary for policy consideration.

Ogundele and Okoruwa (2004) examined technical efficiency differentials between farmers who planted traditional rice varieties and those who planted improved varieties in Nigeria using stochastic production frontier. Results showed that significant increase recorded in output of rice in the country could be traced mainly to area expansion. Other variables that contributed to technical efficiency were; hired labour, herbicides and seeds. The average technical efficiency was 90 and 91 percent for traditional and improved rice variety farmers, respectively. Further analysis showed that farmers in both categories were operating at a point of increasing return to scale. The test of hypothesis on the differentials in technical efficiency between the two groups of farmers showed that there was no absolute differential in technical efficiency between them.

Amaza and Maurice (2005) investigated factors that influence technical efficiency in rice-based production systems among fadama farmers in Adamawa State, Nigeria. A Cobb-Douglas stochastic frontier production function, which incorporates technical inefficiency model, was estimated using the maximum likelihood estimation (MLE) technique. Technical efficiencies vary widely among farms, ranging between 0.26 and 0.97 and a mean technical efficiency of 0.80 implying that efficiency in rice production among fadama farmers in Adamawa State could be increased by 20 percent through better use of available resources, given the current state of technology. The inefficiency model reveals that farming experience and education significantly affect farmers efficiency levels.

Umeh and Asogwa (2005) analyzed the effect of some government policy packages on the technical efficiency of cassava farmers in Benue State, Nigeria. The study used the Cobb-Douglas frontier production function and assumed a truncated normal distribution for the inefficiency term. Cross-sectional data was used. The parameters of the model were estimated by the maximum likelihood estimation method. Their results show that majority (63.6 percent) of the cassava farmers operated close to the frontier production function. The estimated technical efficiency scores varied between 31 percent and 100 percent with a mean score of 89 percent. The findings showed that cassava production in the state can be improved by increasing farmers' access to policy packages such as extension services, market access, improved cassava variety and processing technology.

Ogundari (2006) employed a stochastic frontier profit function to analyse determinants of profit efficiency among scale rice farmers in Nigeria. The obtained mean profit efficiency of 60 percent. The results also showed age education, farming experience and household size has positive and significant effect on profit efficiency.

Ogundari and Ojo (2006) examined the production efficiency of cassava farms in Osun state of Nigeria using farm level data. The stochastic frontier production and cost function model were used to predict the farm level technical and economic efficiencies, respectively. Their results shows that mean TE, EE and AE of 0.903, 0.89 and 0.807 were obtained from the analysis respectively meaning that TE appears to be more significant than AE as a source of gain in EE.

Okoruwa et al. (2006) analysed technical, allocative and economic efficiency of upland and lowland rice producers in Niger State Nigeria using a stochastic production function efficiency decomposition methodology. They obtained an average technical efficiency of 81.6 percent for upland rice and of 76.9 percent for lowland rice. The analysis of variance (ANOVA) was used to investigate the association between EE, TE and AE, and seven socioeconomic characteristics. They found that experience, household size, farm size, sex and improved rice variety has significant impact on rice farmers. Their results showed that farmers could increase output and household income through better use of available resources given the state of technology in terms of improved varieties of rice seeds.

Ogundari et al. (2006) estimated a Cobb-Douglas cost frontier function in order to examine economies of scale and cost efficiencies of small scale maize farmers in Nigeria using a cross-sectional data on 200 farms. The maximum likelihood estimates of the frontier cost function and the inefficiency model were obtained simultaneously in a one-stage procedure. They obtained mean cost efficiency of 1.16 implying that an average maize farm in the area has costs that are 16 percent above the minimum defined by the frontier. About 83 percent of the farms included in the sample operated close to the frontier level. Farming experience and age were found to have significant effect on the cost efficiency of the farmers.

Okoye et al. (2006) employed stochastic frontier translog cost and production functions to measure the level of allocative efficiency and it's determinants in small-holder cocoyam production in Anambra state, Nigeria. The parameters of the stochastic frontier cost function were estimated using the maximum likelihood method. The result of the analysis shows that individual farm level allocative efficiency was about 65 percent. The study found age and education to be negatively and significantly related to allocative efficiency. Farm size coefficient also had a negative relationship with allocative efficiency and was significant. Fertilizer use, credit access and farm experience was significant and directly related to allocative efficiency.

Amos (2007) estimated a stochastic frontier by maximum likelihood method to examine the productivity and technical efficiency of Crustacean production in

Nigeria. A Cobb-Douglas stochastic frontier production function was estimated using primary data. Two models were tested for the presence of technical inefficiency effects using the log likelihood ratio (LR) test. The model without the inefficiency term was dropped. The technical efficiency of producers ranged between 0.45 and 0.98 with a mean of 0.70. The result showed that age and level of education were an increasing function of technical inefficiency while family size and leadership role were decreasing functions of technical inefficiency. Although the sign of the education variable was contrary to the a priori expectation, the explanation given for this is that probably the more educated the producers are, the less time they devote to Crustacean production and the more time they devote to other activities such as politics and merchandising as a form of income diversification. This study also found that cost of fishing equipments and other production costs had significant influence on Crustacean production in Nigeria. The study therefore recommended that producers be encouraged to use better fishing nets and motorized outboard engines to increase their production. It appears that productivity and efficiency were treated as same in this study as there was no evidence of measuring productivity as a separate variable.

Idiong (2007) employed a stochastic frontier production function that incorporated inefficiency factors to provide estimates of technical efficiency and its determinants using data obtained from 112 small scale swamp rice farmers in Cross River State. The results indicated that, the rice farmers were not fully technically efficient. The mean efficiency obtained was 77 percent indicating a 23 percent allowance for improving efficiency. The result also shows that, farmers' educational level, membership of cooperative/farmer association and access to credit significantly influenced the farmers' efficiency positively.

Adewumi and Adebayo (2008) used a cross-sectional data from 152 sweet potato farmers from Kwara State, Nigeria to measure the profitability and technical efficiency of these farmers. For estimation of technical efficiency, they assumed a Cobb-Douglas stochastic frontier production function and estimated the model using maximum likelihood method. A mean technical efficiency score of 0.44 was obtained showing there is considerable inefficiency among the sweet potato farmers. Farm size, education, access to credit, contact with extension agents were found to have significant influence on technical efficiency of the farmers thus improving these variables could increase their technical efficiency.

Ajibefun (2008) assessed the sensitivity of technical efficiency predictions to the choice of estimation method by comparing results from parametric SFPF and non-parametric DEA. The SFPF mean technical efficiency (0.68) was somewhat higher than that from DEA (0.65) and this was explained by the fact that the DEA model being non-stochastic reports noise as inefficiency hence its lower mean technical efficiency. The study also observed dissimilar distributions of efficiency distribution. The study did not indicate if these differences were statistically significant. The study however found that both methods produced similar result for age and education variable with respect to the sign and significance of their impact on efficiency.

Kareem et al. (2008) applied the stochastic frontiers production analysis to estimate the technical, allocative and economic efficiency among the fish farmers using concrete and earthen pond systems in Ogun State. Mean technical efficiency in the concrete pond system was 88percent while earthen pond system was 89 percent. Similarly, the allocative efficiency results revealed that concrete pond system was 79 percent while earthen pond had 85 percent. The results of economic efficiency also revealed an average of 76 percent in concrete pond system while it was 84 percent in the earthen pond system. Further analysis revealed that pond area, quantity of lime used, and number of labour used were significant factors that contributed to the technical efficiency of concrete pond system while pond, quantity of feed and labour are the significant factors in earthen pond system.

Oyekale and Idjesa (2009) employed the stochastic production function to analyse adoption of improved maize seeds and technical efficiency of maize farmers in Rivers State Nigeria. Their results show that use of hybrid seeds, experience, crop rotation, minimum tillage, fertilization and age significantly reduces inefficiency. This study however did not provide estimates of technical efficiency of maize farmers.

Okoruwa et al. (2009) examined the relative economic efficiency of small and large rice farms in North Central, Nigeria. They found that the use of modern rice varieties significantly increases profits. Significant difference in economic efficiency between small and large farms was also discovered. Therefore, it is suggested that, to improve technical efficiency of rice farms, an accelerated program to provide modern rice varieties, fertilizer and land availability is needed. The paper provided support to eliminate bias distribution of production inputs to large rice farms.

Ojo et al. 2009 examined the implication of resource productivity and farm level technical inefficiency in yam production on food security in Niger state, Nigeria using a stochastic frontier production function. Their findings showed the return to scale of 1.686 indicating an increasing return to scale. The study also showed that the levels of technical efficiency ranged from 31.72 percent to 95.10 percent with mean of 75.64 percent which suggests that average yam output falls 24.46 percent short of the maximum possible level. Their result further showed that, farmers' educational level, years of farming experience and access to extension service had significant and positive impact on farmers' efficiency. Thus, it was recommended that relevant policies that would enhance the technical skill of the farmers and access to extension services should be evolved by the stakeholders.

Okoye et al. (2009) employed a Cobb-Douglas stochastic frontier production function to examine the relationship between farm size and technical efficiency in small holder cassava production in Ideato LGA of Imo state using data from a 2008 farm-level survey of 90 rural households. The study showed a strong inverse relationship between farm size and technical efficiency. They concluded that policies of de-emphasizing cassava production in the estate sector while encouraging it in smallholdings will foster equity and efficiency. Therefore, the study recommended land redistribution policies targeted towards giving lands to the small-holder farmers.

To conclude on the Nigerian studies, a detailed review of a meta analysis of technical efficiency in Nigerian agriculture by Ogundari (2009) is provided. A variety of sources were explored to compile the list of papers cited in the study. The analysis was performed with a truncated regression on a total of sixty four studies covering the period 1999-2008. Only studies with the application of primal- stochastic frontier production model were used because studies based on dual representations of the technology frontier as well as non-parametric (e.g., DEA) models in Nigerian agriculture obtained were insignificant in number. None of the studies used panel data showing the dearth of panel survey in Nigeria. The study showed that 63 percent of technical efficiency studies in Nigeria were conducted on food crops showing the dominant position of this sub-sector in Nigerian agriculture. Sixty (69) percent of the studies used single output while 31 percent used aggregate output. Of the studies that employed a single output approach, only one was on maize despite the importance of maize in Nigeria. Cobb-douglas functional form was employed by 88 percent of the studies while only 12 percent employed translog form. Whereas 49 percent of the studies were conducted in the Southeast zone, only 12 percent of such study was conducted in the Northcentral zone (the intended study area for this current study). The results showed that mean technical efficiency (MTE) in Nigerian agriculture increased significantly over the years. The overall average MTE computed from all the studies was 0.739 which is significantly not different from 0.737 obtained by Bravo-Ureta et al. (2007) for African countries and 0.68 obtained by Thiam et al. (2001) for developing countries. This finding, however, suggests that, there is a large potential for improvement in Nigerian agricultural production systems, as about 26 percent of the agricultural output in the country could be expanded without any additional use of inputs in comparison to what could be achieved under full technical efficiency. The findings further showed that studies in the Southwest region of the country produced higher MTE with average of 0.842 whereas the average is 0.720 for Northcentral zone implying that improving efficiency and productivity in Nigerian agriculture might require regional specific-policy responses. Regarding the unconditional effect of the choice of functional form, the study observed an average MTE of 0.79 for studies with Cobb-Douglas, and 0.69 for translog. In contrast, Thiam et al. (2001) and Bravo-Ureta et al. (2007) reported higher average MTE for studies with translog compared to Cobb-Douglas. The study however, found no statistical difference between the MTE of Cobb-Douglas and translog in this study. Study specific-characteristics such as sample size, number of inputs used as well as studies with focus on crop and livestock production were found to significantly impact MTE. Within the sample, seventy one observations contain quantitative results on sources of technical efficiency differences usually incorporating socio-economic variables. Based on this, fifty three percent identified educations as a significant determinant of technical efficiency while thirty eight percent showed that experience is important. Extension was shown to be an important determinant by twenty three percent of the observations while nineteen percent identified age as significant determinant of technical efficiency in Nigerian agriculture over the years.

All the above studies that compared distance function and other approaches limited their analysis to technical or environmental efficiency only. This is not surprising given the methodological complexity involved in decomposing cost efficiency into its technical and allocative components. Further, with exception of the Herrero (2005) and Cuesta et al. (2009) studies, all others considered parametric deterministic distance functions, thus the possibility of stochastic noise in the data was ignored. Given that the focus of this study is on agriculture which is well known to be affected by factors such as weather and macro economic factors that are beyond the control of farmers, the neglect of random factors may have serious implications on the study conclusions. Thus, the current study intends to fill these gaps by making a comparison of parametric stochastic input distance functions (SIDF), non-parametric data envelopment analysis (DEA) and conventional parametric stochastic frontier production frontier (SFPF) approaches to analysis of technical, allocative and cost efficiency and their determinants in the Nigerian maize sector. Based on the result of the comparative analysis an integrated model is developed for resolution of model selection difficulties.