CHAPTER-4

METHODOLOGY
This chapter deals with the methodological developments and techniques adopted for data analysis. Since, the secondary data from Cost of Cultivation Schemes of Directorate of Economics and Statistics has been used in this study, the sampling design and data has also been explained. For convenience of explanation, this chapter is divided into following two sections.

- Sampling Design and Data
- Analytical Framework

4.1 SAMPLING DESIGN AND DATA

The source of secondary data is “Cost of Cultivation of Principle Crops” scheme Directorate of Economic and Statistic. The data at farmer’s level is collected by the state Agriculture University of the respective States on a schedule (questionnaire) and send to the Directorate of Economic and Statistic, Department of Agricultural and Co-operation, under Ministry of Agriculture, government of India at Shastri Bhawan New Delhi. Here the data is compile in the package develop by FAO know as “Farm map data management system”. The data is entered in the tabulation sheets in a coded form for the certain year. The data is analyzed by this department for predicting agricultural policies at macro-level, declaring support prices for the major crops etc\(^1\).

The design followed by them is “Crop cutting Experiments” for the Yield estimates. A three stage stratified random sampling method is used for the selection,

\(^1\) DES is the central co-ordinating agency, which is responsible for the collection, compilation and publication of agricultural statistics at all-India level.
with Tehsils (each contains 100 to 300 villages) as first stage unit or strata. A village/cluster of village as the second stage units and holding as the third and ultimate sampling unit. Each state is demarcated into homogeneous agro-climatic zones based on cropping pattern, soil type, rainfall, etc. The primary sampling units (tehsils) are allocated to different zones in proportion to total area of all crops covered in the study. The primary sampling units are selected in each zone (stratum) with probability proportional to the area under the selected crops, and with replacement. Within each tehsil, one village having more than 200 optional holdings is selected and defined as nucleus village. If the number of operational holding is less than 200 in the nucleus village another village located at south direction is selected, if required the third village may also be selected to make the number of operational holding 200. In each selected village/cluster, all the operational holdings are enumerated and classified according to size into 5 size classes, the class limits being fixed uniformly for all village/clusters as:

<table>
<thead>
<tr>
<th>Group</th>
<th>Size of holding</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.99 ha.</td>
</tr>
<tr>
<td>2</td>
<td>1-1.99 ha.</td>
</tr>
<tr>
<td>3</td>
<td>2-3.99 ha.</td>
</tr>
<tr>
<td>4</td>
<td>4-4.99 ha.</td>
</tr>
<tr>
<td>5</td>
<td>6 ha. &amp; above</td>
</tr>
</tbody>
</table>

In each size class, two holdings are selected by simple random sampling without replacement. Only 10 farmers are selected out of 200 operational holdings. Sample farmers are changed after three years.

For our present study the state of Haryana is chosen. In Haryana data is collected from the Schedule (questionnaire) surveyed by Chaudry Charan Singh Haryana Agriculture University (CCSHAU) at Hissar. The crop chosen for the study purpose are wheat and rice as they are the two major food crops of the State. Two major wheat-rice growing districts are Karnal and Kurukhetra and hence considered for study. The panel
data for the period 1985–1995 of 10 farmers each from two districts is considered. The data was classified into two periods VIZ; 1985-1990, 1990-1995 and 1985-1995 for functional analysis and denoted as Technology-1, Technology-2 and pooled data as overall, respectively. This is done on the sound logical base that the technological adoption and productivity was on accelerating trend during period 1985-90 and became stagnant during 1990-95 because of environmental degradation accompanied together with slow pace of technological innovations in agriculture.

4.2 ANALYTICAL FRAMEWORK

In this chapter, the analytical framework to be used for the analysis of technical efficiency is discussed in detail under the broad spectra of following two headings:

- Estimation of the model
- Specification of the model

4.2.1 Estimation of the Technology Frontier Model

Technology frontier varies with regards to the assumptions made on the outer bound of the Frontier, which may be either Deterministic or Stochastic, and with regard to the measurement approach, which may be either Non-Parametric or Parametric. Now one of the most important question remained to be answered is – what would be the appropriate distributional assumption for (error term) outer bound of the frontier and what would be the estimation technique in order to give most reliable estimate of the technology frontier and ensuring thereby correct measure of efficiency.

Production technology is represented by the transformation (production) function that defines the maximum attainable outputs from different combinations of inputs.
Hence, the transformation function describes a boundary or a frontier. If the production function frontier is known, the technical inefficiency of any firm can be assessed easily by simply comparing the position of the firm relative to the frontier. There are two competing paradigms on how to construct frontiers. One use mathematical programming techniques, the other employs econometric techniques. The chief advantage of mathematical programming approach is that no explicit functional form needed to be imposed on the data. However, the calculated frontier may be warped if the data are contaminated by statistical noise. The econometric approach can handle noise, but it imposes an explicit, and possibility overly restrictive, functional form for technology. Numerous methods have been developed for the empirical measurement of frontier functions and the potential deviations from such functions these methods can be categorized under the broad heading as follows:

- Parametric Approach of Estimation of Production Frontier
- Non-Parametric Approach of Estimation of Production Frontier

4.2.2 Parametric Approach of Estimation of Technology Frontier Model

There are mainly two approaches for measurement of technical efficiency through frontier functions which will be tested in the present study.

- Deterministic Technology Frontier Model
- Stochastic Technology Frontier Model

4.2.3 Estimation of Deterministic Technology Frontier Model

The frontier is called deterministic with the expected notion that all the observation lie exactly on or below production frontier. This is also referred as full frontier in the literature. Here the error term is totally attributed to the technical efficiency of the farm or firm. Deterministic frontier directly yields estimates of individual firms
efficiency as the residuals from estimation. Let us consider a production relationship (function) of the following type;

\[ Y_t = \alpha + \beta X_t + \varepsilon_t \]

\( t = 1,2, \ldots \ldots 10 \)

Where;

- \( Y_t \) is the vector of observation on outputs
- \( X_t \) is the matrix of observations on inputs
- \( \alpha \) is the intercept term
- \( \beta \) is the vector of coefficients of production function.
- \( \varepsilon_t \) is the vector of inefficiency component i.e. shortfall of actual output compare to potential level.

In order to arrive at estimates of the frontier function, the error distribution must be skewed with negative expectations \( \{ E(\varepsilon) < 0 \} \), reflecting the presence of technical inefficiency in the production process.

There are two major statistical techniques involved in estimation of Deterministic Technology Frontier. The choice of distributional forms for the error plays a key role in application of both of the techniques. These two techniques are:

i) Maximum Likelihood Estimation Technique (MLE).

ii) Corrected Ordinary Least Square Technique (COLS)

### 4.2.4 Maximum Likelihood Estimation Technique (MLE)

The Maximum Likelihood Estimation is a method, which on fundamental way largely depend on the type of specification made regarding distributional forms used in
formulation of likelihood function. It is because of this reason different assumed distributions lead to different estimates of the frontier models. Schmidt (1976) observed that the Agner-Chu optimization criteria could be interpreted as log likelihood function for models in which one sided residuals were distributed as exponential and half normal. However, a particular problem arises in that the gradients of both log likelihood have non-zero expectations, and the Hessian of both log likelihood function are singular. This was because of the violation of one of the familiar regularity conditions, that the range of the observed variable $y_j$ is dependent upon the parameters to be estimated. Greene (1980) gave solution to the problem and suggested that the usual desirable properties of maximum likelihood estimates still holds if the density of the distribution of $\varepsilon$ satisfies the following conditions:

(a) The density of $\varepsilon$ is zero at $\varepsilon$

(b) The derivatives of the density of $\varepsilon$ with respect to its parameter approaches zero as $\varepsilon$ approaches zero.

Whole family of Pearson's distributions follows the above conditions and hence could be used. Four distributions for estimation of Deterministic Frontier are assumed in the literature-

4.2.4.1 Gamma Distribution

This distribution is extremely attractive for the estimation of Frontiers as it implies that the model is quite flexible in the shape of the error distribution it will accommodate. The regularity condition for Maximum Likelihood Estimation are satisfied easily by the form

The two parameter gamma density function is written as

$$f(\varepsilon) = G(\lambda, p) = \frac{\lambda^p \varepsilon^{p-1} e^{-\lambda \varepsilon}}{\Gamma(p)}$$
where \( \epsilon \geq 0, \lambda > 0, p \geq 2 \)

\[
\text{Mean} = E(\epsilon) = \frac{p}{\lambda} \quad \text{& Variance} = V(\epsilon) = \frac{p}{\lambda^2}
\]

The log of the likelihood function for this disturbance model is

\[
\log L = TP \log \lambda - T \log \left( \sqrt[p]{(p)} + (p-1) \sum \log (y_t - \alpha - \beta X_t) - \lambda \sum (y_t - \alpha - \beta' X_t) \right) \quad (5)
\]

The first derivatives of the log likelihood are:

\[
\begin{align*}
\frac{\partial \log L}{\partial \lambda} \\
\frac{\partial \log L}{\partial p} \\
\frac{\partial \log L}{\partial \alpha} \\
\frac{\partial \log L}{\partial \beta}
\end{align*}
\]

\[
\begin{bmatrix}
TP/\lambda - \Sigma e_t \\
T (\log \lambda - \sqrt[p]{(p)} + \Sigma t \log \epsilon_t) \\
T\lambda - (p-1) \Sigma (1/\epsilon_t) \\
T\lambda X' - (p-1) \Sigma (1/\epsilon_t) X_t
\end{bmatrix}
\quad (6)
\]

Now we see that

\[
1 \text{ restriction } p > 2 \text{ is necessary to get following condition}
\]

a) \( f(y, \phi)'y = |\phi| = f_{\epsilon}(0) = 0 \)

b) \( [\partial f(\epsilon)/\partial \epsilon]_{\epsilon = 0} = 0 \)
Where \( \phi = (\alpha, \beta, \gamma) \)

Which is required condition to prove Crammer Rao inequality

Hence equation (6) gives

\[
E \left( \frac{\partial \log L}{\partial \phi} \right) = E \left( \frac{\partial \log L}{\partial \alpha} \right) = E \left( \frac{\partial \log L}{\partial \beta_k} \right) = 0
\]

We get estimate of \( \lambda, \alpha, \beta_k \) and since they are MLE estimates hence they possess all the desirable properties for best estimate except unbiasedness.

4.2.4.2 **Half normal Distribution**

The density function of half normal distribution is given by

\[
f_\varepsilon(\varepsilon) = \frac{2}{\theta \sqrt{\pi}} e^{-\varepsilon^2/2\theta^2}, \quad \varepsilon \geq 0, \theta > 0
\]

\[
E(\varepsilon) = \theta \sqrt{2}/\sqrt{\pi}
\]

\[
V(\varepsilon) = \theta^2 (\pi - 2)/\pi
\]

The log of the likelihood function are obtained by

\[
\left( \frac{\partial \log L}{\partial \phi} \right) = 0 \quad \text{where } \phi = (\phi, \alpha, \beta)
\]
Now here we see that

\[
E (\partial \log L / T \partial \phi) = \begin{pmatrix}
0 \\
\sqrt{2}/\theta \sqrt{\pi} \\
\sqrt{2X}/\theta \sqrt{\pi}
\end{pmatrix} \neq 0
\]

This result violate one of regularity conditions used to prove that t MLE are consistent, efficient and asymptotically normal. This is because the matrix obtained is singular and we can not estimate the desired coefficients.

### 4.2.4.3 Exponential distribution

The density function of exponential distribution is given by

\[
f(\varepsilon) = \lambda e^{\lambda \varepsilon}, \varepsilon \geq 0
\]

\[
E(\varepsilon) = (-1/\lambda) & V(\varepsilon) = 1/\lambda^2
\]

The derivatives of log likelihood function are obtained by equating

\[
(\partial \log L / \partial \phi) = 0 ; \text{ where } \phi = (\lambda, \alpha, \beta)
\]

Now we see

\[
E = (1 \partial \log L / T \partial \phi) = \begin{pmatrix}
0 \\
\lambda \\
\lambda x
\end{pmatrix} \neq 0
\]
This result violate one of regularity conditions used to prove that t MLE are consistent, efficient and asymptotically normal. This is because the matrix obtained is singular and we can not estimate the desired coefficients.

For overcoming the above difficulty in both the distribution we can either use distribution having density

\[ f_\epsilon(0) = 0 \text{ or construct a more reasonable error structure than a purely one-side one i.e.} \]

\[
\epsilon_i / \sqrt{1 - \theta} \quad \text{if } \epsilon_i^* > 0, \quad t = 1, \ldots, T \\
\epsilon_i = \frac{\epsilon_i^*}{\sqrt{\theta}} \quad \text{if } \epsilon_i \leq 0
\]

Where \( \epsilon_i^* \sim N(0, \sigma^2) \) if \( 0 < \theta < 1 \), otherwise \( \epsilon_i^* \) has either the negative or positive truncated normal distribution, when \( \theta = 1 \) or \( \theta = 0 \) respectively.

### 4.2.4.4 Truncated Normal Model

The probability density function of random variable \( u \) that is normally distributed with mean \( \mu \) and variance \( \sigma^2 \) and truncated from below at a point \( \alpha \) is

\[
f(u/u > \alpha) = \begin{cases} 
  \frac{f(u)}{1 - \eta(\alpha - \mu)} \\
  \frac{f(u)}{\sigma}
\end{cases}
\]

Where \( \eta(.) \) is the CDF of a SNV and \( f(u) \) is the PDF of the truncated normal random variable \( u \).
it's mean and variance are,

\[
E\{x/x > \alpha\} = \begin{pmatrix}
\mu + \sigma \cdot \phi(\alpha - \mu) \\
\frac{1}{\sigma}
\end{pmatrix}
\begin{pmatrix}
\frac{1}{\sigma} \eta(\alpha + \mu) \\
\sigma
\end{pmatrix}
\]

and

\[
\text{Var}\{x/x > \alpha\} = \sigma^2(1 - \delta(a))
\]

respectively,

where

\[
\delta(a) = \left( \frac{\eta(\alpha - \mu)}{\sigma} \right) \left( \frac{\eta(\alpha - \mu)}{\sigma} \right) - \left( \frac{\alpha - \mu}{\sigma} \right)
\]

4.2.5 Corrected Ordinary Least Square Technique

Here except nonzero mean and non-normality all the usual assumptions of OLS theory are followed. The estimation of production elasticities are not effected in this specification as the deterministic frontier comes about by a mere upward shift on the scale of corresponding 'average' production function. This technique is called Corrected Ordinary Least Square because we correct the constant term of the obtained OLS technique by using the parameters of different distributions to get the frontier function. Hence we get different estimates by using different distributions and the appropriateness of using one over other is also questionable. Here there are two ways by which we can obtained the estimated coefficients. In the method given by Kopp and Smith (1980) the equation of production function is first solved by using the ordinary least square technique and then the intercept term is shifted up until all residuals are positive and non
zero i.e., we only shift the intercept of the estimated function until all residual have the correct sign. The corrected constant term is \(a + \max(\varepsilon_i)\). In the second approach, different distributions are used to obtain the coefficients of the model i.e., we first estimate the mean and variance of the required distribution and then add this variance to the intercept term and then again estimate the coefficients of the model. Here one thing is to be noted that correction to the constant term is not independent of the distribution assumed hence different distribution yield different results. The distribution used in COLS technique are:

### 4.2.5.1 Gamma Distribution

The two parameter gamma density function is written as

\[
f(\varepsilon) = G(\lambda, \theta) = \frac{\lambda^\theta \varepsilon^{\theta-1} e^{-\lambda \varepsilon}}{\Gamma(\theta)}
\]

where \(\varepsilon > 0, \lambda > 0, \theta \geq 2\)

where \(E(\varepsilon) = \lambda\) & \(V(\varepsilon) = \lambda\). Now the equation (3) changes to

\[
Y = (\alpha - \lambda) + \beta X_\varepsilon (\varepsilon - \lambda)------------------------------(4)
\]

We regress this equation again to obtain estimates of technical inefficiency.

### 4.2.5.2 Exponential distribution

The density function of exponential distribution is given by

\[
f(\varepsilon) = \lambda e^{-\lambda \varepsilon}, \varepsilon \geq 0
\]

where

\[
E(\varepsilon) = (\lambda) & V(\varepsilon) = (\lambda^2), \text{ hence equation (3) changes to}
\]
We regress this equation again to get the estimates of technical inefficiency.

4.2.5.3 Half Normal distribution

The density function of half normal distribution is given by

$$f_\varepsilon(\varepsilon) = \frac{2}{\sqrt{\theta\pi}} e^{-\varepsilon^2/2\theta^2}$$

where $\varepsilon \geq 0$, $\theta > 0$

$$E(\varepsilon) = \theta \sqrt{2/\sqrt{\pi}}$$ and $$V(\varepsilon) = \theta^2 (\pi - 2/\sqrt{\pi})$$ and the equation becomes

$$Y = (\alpha - \theta^2) + \beta X_r (\varepsilon - \theta^2)$$

The equation (4), (5) and (6) are solved by using OLS technique. Truncated normal and half normal distribution provide $\sigma$ as the corrected term. The coefficients obtained using different forms are different and also with data sets different provide different results and hence different technical inefficiencies, except for the special case when $V(\varepsilon) = 1$.

4.2.5.4 Truncated Normal Model

The probability density function of random variable $u$ that is normally distributed with mean $\mu$ and variance $\sigma^2$ and truncated from below at a point $\alpha$ is
Where \( \eta(.) \) is the CDF of a SNV and \( f(u) \) is the PDF of the truncated normal random variable \( u \).

It’s mean and variance are,

\[
E[x/x>\alpha] = \frac{\mu + \sigma \cdot \phi(\alpha - \mu)}{1-\eta(\alpha + \mu)}
\]

and

\[
Var[x/x>\alpha] = \sigma^2(1-\delta(a))
\]

where

\[
\delta(a) = \frac{\eta(\alpha - \mu)}{\sigma} \cdot \frac{\eta(\alpha - \mu)}{\sigma} - \frac{(\alpha - \mu)}{\sigma}
\]
4.2.6 Estimation of Stochastic Technology Frontier Model

The notion of deterministic frontier shared by all firms ignores the very real possibility that a firm's performance may be control by the factors not in its control (such as poor machine performance, bad weather, input supply break downs and so on), as well as factor under its control (inefficiency). The above notion lies behind the Stochastic Frontier (also called "composed error model"\(^1\)). The essential 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 firm's control is unavoidable. The Stochastic Frontier Model checks the overestimation of measures of efficiency and gives the correct estimation for which the necessary action may be taken.

This additional flexibility that this model provided was the variation of actual output below the production frontier no longer needs to be ascribed solely to technical inefficiency and can instead be divided between technical inefficiency and random variations in the production conditions over which the producer has little control.

A Stochastic production frontier model may be written as

\[
Y = f(X_{it}; \beta) \exp(u-v) \tag{6}
\]

Where \(f(X_{it}; \beta)\) is the deterministic part of the production function

\(v\) is a random disturbance

\(u\) is a non-negative one-sided distribution term, capturing technical inefficiency.

In log form it is written as

\[
\ln Y = \ln f(x) + v - u \tag{7}
\]

\(^1\) Proposed by Aigner et.al (1977) and worked by Meeusen and Van Den Broeck (1977).
Here $u$ can also be interpreted as the distance of actual production from the production frontier. The negative sign merely reinforces the idea of the shortfall from the maximum. Direct estimates of the stochastic production frontier model may be obtained by applying the following two techniques:

i) Maximum Likelihood Estimation Technique (MLE).

ii) Corrected Ordinary Least Square Technique (COLS).

### 4.2.7 Maximum Likelihood Estimation Technique

Introducing specific probability distributions for $v$ and $u$, assuming that $u$ & $v$ are independent and that $x$ is exogenous. The asymptotic properties of the maximum likelihood estimators can be proved in the usual manner. The presence of the symmetric error component $v$ solves the bounded range problem encountered by some variants of the deterministic frontier model.

Four distributions for the one-sided error term have been assumed in the literature: half normal, truncated normal, exponential and gamma, with the two sided disturbance assumed to be normally distributed with zero mean. Aigner, Lovell & Schmidt (1977) discussed two distributions, the half normal & the exponential for the one sided error term representing technical inefficiency while Stevenson (1980) first suggested the truncated normal & Greene (1980, 1990, 1993) has advocated two parameter gamma distribution. They are as follows:

i) MLE of NHNM (normal half normal model)

ii) MLE of NTNM (normal truncated normal model)

iii) MLE of NEM (normal exponential model)

iv) MLE of NGM (normal gamma model)
The probability density function of a random variable $u$ that has a half-normal distribution is

$$
\frac{1/2}{\pi} \left( \frac{1}{\sigma^2} \right)^{1/2} \exp \left( -\frac{u^2}{2\sigma^2} \right)
$$

and its mean and variance are

$$
E(u) = \frac{2^{1/2}}{\pi} \sigma u
$$

$$
V(u) = \frac{(\pi-2)}{\pi} \sigma^2 u
$$

respectively.

The probability density function of a random variable $(v-u)$ where $v$ is a normal random variable with mean 0 and variance $\sigma^2_v$ is given by

$$
f(v-u) = \frac{1/2}{\pi} \exp \left( -\frac{(v-u)^2}{2\sigma^2} \right) \times \left[ 1 - \phi \left( \frac{v-u}{\sigma} \right) \right]
$$

Where $\phi(.)$ is CDF of the standard normal distribution

if $\sigma^2 = \sigma_u^2 + \sigma_v^2$.

$\lambda = \sigma_u/\sigma_v$
and given that the Jacobian of the transformation from the unobserved disturbances to the observed regressors is 1, the log of the likelihood function is given by

\[
\log L = \sum_{t=1}^{T} \log f(V_t - U_t) = \sum_{t=1}^{T} \log f(\varepsilon_t)
\]

\[
= \sum_{t=1}^{T} \left( -\frac{1}{2} \log(2\pi) + \log(\sigma^2) - \frac{\varepsilon_t^2}{2} + \log \left[ 1 - \frac{\varepsilon_t \phi(\varepsilon_t)}{\sigma} \right] \right)
\]

Maximising the likelihood function with respect to the observed variables yields efficient estimates of the parameters of the models. But this would yield only an estimate of sample or aggregate technical efficiency.

Jondrow, Lovell, Materov and Schmidt (1992) and Kalirajan and Flinn (1983) provided a procedure for a farm level estimates which uses distribution of the one-sided error term conditional on the estimated of the composite error term. Both the mean and the mode of this conditional distribution may be used to estimate observation-specific value of \( u \). Specifically, if \( u \) is assumed to have a half-normal distribution, the conditional distribution of \( u \) given \( \varepsilon \) is that of a \( N(\mu^*, \sigma^*) \) random variable truncated at zero.

We can use either mode of this conditional distribution

\[
M(u|\varepsilon) = -\varepsilon \begin{cases} [\sigma^2_u] & \text{if } \varepsilon = 0 \\ \sigma^2 & \text{if } \varepsilon > 0 \end{cases}
\]

\( ^1 \) Before this, the problem of deriving firm specific estimates of \( u \) from the average sample value of \( u \) was major drawback of stochastic frontier approach.
or the mean of this conditional distribution

\[
E [u_i/\epsilon_i] = \frac{\eta(\epsilon_i \lambda_i)}{1 - \phi(\epsilon_i \lambda/\sigma)} - \frac{\epsilon_i \lambda_i}{\sigma} \\
= \frac{\sigma \lambda}{(1 + \lambda^2)} \begin{pmatrix} \eta(\epsilon_i \lambda) & (\epsilon_i \lambda) \\ \phi(-\epsilon_i \lambda) & \sigma \end{pmatrix}
\]

Where \( \sigma^2 = \sigma_v^2 + \sigma_u^2 \)

\[
\mu^* = \frac{-\sigma^2 (v-u)}{\sigma^2}
\]

\[
\lambda = \frac{\sigma_u}{\sigma_v}
\]

\[
\sigma^* = \frac{\sigma_v^2}{\sigma_u^2}/\sigma^2
\]

and \( \eta(\cdot) \) and \( \phi(\cdot) \) are respectively the probability density function (PDF) and cumulative distribution function (CDF) of the standard normal distribution, evaluated at \( \epsilon \lambda/\sigma \).

The above procedure provides observation-specific estimate of \( u_i \). However, although the resulting estimates are unbiased, they are inconsistent\(^2\).

The estimated variance of the one-sided distribution can be calculated as per the definition above, using the estimated value of \( \sigma_u^2 \), i.e. \( S_u^2 \). That is,
Est. $V(u^2) = \frac{\pi - 2}{\pi} S_u^2$  

the estimated variance of the composite error term $S_e^1$ is given by

$$S_e^2 = S_v^2 + S_u^2 = \left( \frac{\pi - 2}{\pi} \right)$$

is simply the sum of the estimated variance of the one-sided error term and that of the stochastic error term.

Calculating the ratio of the estimated variance of the one-sided residuals with respect to that of composite residuals indicates the relative dominance of technical inefficiency over stochastic variation

$$\text{Est}[\lambda] = \left( \frac{\pi - 2}{\pi} \right) S_u^2/S_v^2$$

The above expression gives the estimated value of $\lambda$.

4.2.7.2 The Normal - Truncated Normal Model

The probability density function of random variable $u$ that is normally distributed with mean $\mu$ and variance $\sigma^2$ and truncated from below at a point $\alpha$ is

$$f(u/\alpha) = \begin{cases} \frac{f(u)}{1-\eta(\alpha-\mu)} & \text{if } u > \alpha \\ \frac{f(u)}{\sigma} & \text{otherwise} \end{cases}$$

\[1 \text{ i.e. the variance of the estimate does not reduce to zero with increases in the sample size.}\]
Where \( \eta(.) \) is the CDF of a SNV and \( f(u) \) is the PDF of the truncated normal random variable \( u \).

It's mean and variance are,

\[
E[x/x>a] = \left( \frac{\eta(\alpha - \mu) + a \cdot (a - \mu)}{\eta(\alpha + \mu)} \right)
\]

and

\[
\text{Var}[x/x>a] = \sigma^2 (1-\delta(a))
\]

where

\[
\delta(a) = \frac{\eta(\alpha - \mu)}{\sigma} \left( \frac{\eta(\alpha - \mu)}{\sigma} - \frac{(\alpha - \mu)}{\sigma} \right)
\]

Now the probability density function of the random variable \((v-u)\) in this case is given by

\[
F(v-u) = 2\pi \sigma^{1/2} \exp \left( -\frac{1}{2} \left( \frac{v-u - \mu}{\sigma} \right)^2 \right) \cdot \eta \left( \frac{-\mu}{\sigma \lambda} - \frac{\varepsilon \lambda}{\sigma} \right) \cdot \frac{\lambda}{\sigma} \left( 1 + \lambda^2 \right)^{1/2} \, dv
\]
The log-likelihood function is written as

\[
\text{Log } L = -\frac{1}{2} \log (2\pi) + \log (\sigma^2) + (\varepsilon \lambda)^2/\sigma + \log \eta - \left[ \frac{1}{\sigma (\varepsilon \lambda)/\sigma + \mu \lambda \sigma} \right] - \log \eta(\mu/\sigma)(1+\lambda^2)^{1/2}/\lambda
\]

The estimate of sample-wide technical efficiency that results from the maximization of this likelihood function with respect to the observed data variables can be decomposed into observation specific estimates using a variant of the Jondrow et al 9!982) procedure outline above. Specifically, the mean of the conditional distribution is

\[
\text{E}[\varepsilon_i/\epsilon_i] = \left( \frac{\sigma \lambda}{1+\lambda^2} \right) \left[ \phi(\epsilon_i \lambda/\sigma + \mu/\sigma \lambda) - \frac{\phi(-\epsilon_i \lambda/\sigma + \mu/\sigma \lambda)}{\phi(-\epsilon_i \lambda/\sigma + \mu/\sigma \lambda)} \right]
\]

Again the mean and variance of \( u \) and \( \varepsilon \) can be calculated and compared.

### 4.2.7.3 Normal-exponential model

An exponentially distributed random variable \( u \) has a probability density function of the form

\[
f(u) = \theta \exp(-\theta u)
\]

and its mean and variances are

\[
E(u) = \theta \quad \text{and} \quad V(u) = \theta^2 \text{ respectively.}
\]

The probability density function of the random variable \( (\nu - u) \) where \( \nu \) is distributed with mean zero and variance \( \sigma^2 \) is

\[
f(\nu - u) = \theta \exp((\nu - u)^2/2) * \left[ \phi(-\nu/\sigma \sqrt{2}) \right] du
\]
and the corresponding log likelihood function is

$$\log L = \sum_{t=1}^{T} \log f(v-u)_t$$

$$= \sum_{t=1}^{T} \left[ \log (\theta + \sigma^2 \theta^2/2 + \epsilon\theta) + \log \left[ \phi((-\epsilon+\sigma^2\theta)/\sigma) \right] \right]$$

the decomposition of the estimated sample technical inefficiency into those of the constituent firms is, as Greene (1993) shows, a direct extension of the result of Jondrow et al., (1982). The mean of the conditional distribution thus is

$$E[u_i/e_i] = [\epsilon_i - \Delta \sigma_i^2] + \sigma_i \left[ \phi(\epsilon_i - \Delta \sigma_i^2)/\sigma_i \right]$$

As above, the estimates of the variance of the individual and composite residuals can be computed from the data and compared.

4.2.7.4 Normal Gamma Model

The probability density function of a random variable $u$ which has a gamma distribution is

$$f(u) = G(\lambda, p) = \theta^p u^{p-1} e^{-\theta u}$$

$$\Gamma p$$

where $P > 0, \theta > 0$, 

110
and it has a mean and variance given by

\[ E(u) = \frac{\mu}{\omega} \quad \text{and} \quad V(u) = \frac{\mu^2}{\omega^2} \]

respectively. The probability density function of \((u-v)\) can be obtained as

\[
f(v-u) = \frac{1}{2\sigma^2} \exp\left(\frac{-1}{2\sigma^2} \left( v - \mu + \sigma^2 \right) \right) \int_{-\infty}^{\infty} \left( u + \sigma^2 \right) \exp\left[ -\frac{1}{2\sigma^2} \left( u + \sigma^2 \right) \right] \, du
\]

The log-likelihood function can be written as

\[
\log L = \sum_{t=1}^{T} \left[ x_{t} \log \theta - \log \Gamma(p) + \epsilon \theta + \theta^2 \sigma^2/2 + \log \left( \phi - (\epsilon - \theta \sigma^2)/\sigma \right) + \log[E(x_{t}^p | x_{t} > 0, \epsilon)] \right]
\]

where the last term is to be numerically approximated. The resulting estimate of average technical inefficiency can be decomposed into farm-specific estimates using the procedure outlined in Greene (1990), which is a variant of the procedure suggested by Jondrow et al. (1982). Thus

\[
E(\frac{u}{\epsilon}) = \int u f(u/\epsilon) \, du = \frac{E[u^p | u > 0, \epsilon]}{E[u^{p-1} | u > 0, \epsilon]} = h(p, \epsilon)
\]

\[
E[\frac{u}{\epsilon}] = h(p-1, \epsilon)
\]

4.2.8 Corrected Ordinary Least Square Technique

The function can be estimated applying COLS technique as discussed for deterministic frontier model. The parameters of the stochastic frontier model can be estimated using the second and third central moments of the OLS residuals, \(m_2\) and \(m_3\). The following distribution are described in literature. These estimators are consistent but inefficient in comparison to MLEs.
4.2.8.1 Half normal model

The moments of the equation of this model are

\[ m_2 = \left[ \pi - \frac{2}{\pi} \right] \sigma_u^2 + \sigma_v^2 \]

\[ m_3 = \left( \frac{\sqrt{2}/\pi}{1 - (4/\pi)} \right) \sigma_u^3 \]

The adjustment of the OLS constant term is

\[ a = \alpha + \sigma_u \left( \frac{\sqrt{2}/\pi}{1 - (4/\pi)} \right)^{1/3} \]

The decomposition of the estimated average technical inefficiency into farm-specific estimates is once again using the Jondrow et al., (1982) procedure.

4.2.8.2 Exponential Model

The moments of the equation of this model are

\[ m_2 = \sigma_v^2 + \frac{1}{\theta^2} \]

\[ m_3 = -\frac{2}{\theta^3} \]

The adjustment of the OLS constant term is

\[ a = \alpha + \frac{1}{(-2/ m_3)^{1/3}} \]

The decomposition of the estimated average technical inefficiency into farm-specific estimates is once again using the Jondrow et al., (1982) procedure.
4.2.8.3 Gamma- normal model

The moment equations for the gamma mode provide the following consistent estimators of the parameters of the model

\[ \hat{\theta} = -3 \frac{m_3}{m_4 - 3 m_2^2} \]

\[ P = - \frac{\theta^3 m_3}{2} \]

\[ \sigma_v^2 = m_2 \frac{-P}{\theta^2} \]

The adjustment of the OLS constant term is

\[ a = \alpha \frac{-P}{\theta^2} \]

The decomposition of the estimated average technical inefficiency into farm-specific estimates is once again using the Jondrow et al., (1982) procedure.

4.2.9 NON PARAMETRIC APPROACH OF ESTIMATION OF TECHNOLOGY FRONTIER

The Fare, Grosskopf and Lovell(1985) approach begins by specifying a transformation function, \( T \), which assumes constant return to scale and strong input disposability:
\[ T = (x, y) : \lambda z \leq x, y \leq \lambda z, \lambda \in R^k_+ \quad (1) \]

Where

\[ x = a (n \times 1) \text{ vector of input}, \]
\[ y = a (m \times 1) \text{ vector of outputs}, \]
\[ k = \text{the number of farms}, \]
\[ X = \text{the (n x k) matrix of observed inputs and} \]
\[ Y = \text{the corresponding (m x k) matrix of outputs}. \]

The intensity with which any activity \((x_i, y_i)\) is utilized is given by the \((k \times 1)\) vector \(z\). The transformation set is illustrated in figure-1 using three farm observation, A, B & C. Assuming constant return to scale, movement along the horizontal axis represent equi-proportional increases in all inputs. The best practice constant-return-to-scale frontier denoted by the ray OT is constructed by using farm B as the reference point. The transformation set is bonded by OT and the horizontal axis.

Individual firm efficiency is determined relative to the constructed technology frontier. For farm A, the maximum potential output, given its observed inputs use \(x_A\) is \(y_A^*\). The over measure of technical efficiency is equal to the ratio of actual to the potential or efficient output. For farm A, this is \(y_A / y_A^*\) which is equivalent to the measure of inefficiency defined by Farrell (1957). The over all measure of technical efficiency for an individual observation \(i\), \(TE_i\), can be expressed as:

\[ TE(x_i, y_i) = \min (\lambda i : (x_i, y_i / \lambda i) \in T) \quad (2) \]

Can be calculated through solution of the following LP problem:

\[ \sum_{i=1}^{n} x_{ij} z_i \geq x_i, \quad j = 1,2,\ldots,n \quad (3) \]
The optimum input usage is determined as a weighing of the actual input use for all farms. For each input, the theoretically efficient farm will employ an amount that is less than or equal to the quantity used by farm $k$ to produce its current level of output $y_k$.

Figure 1: Transformation set (T) under constant returns to scale

The output constraint consists of two parts. The component $y_1 z_1 + y_2 z_2 + \ldots + y_n z_n$, represents the maximum output of a hypothetically efficient farm, given the actual level of input used by the $i$th farm. Subtracted from this component is the level of output of farm $i$ divided by its level of inefficiency $\lambda$. In Figure 1, output of farm B is the same as the output of the theoretically efficiency farm and consequently $\lambda_B = 1$ in order for potential and actual output to be equal. In contrast, maximum output is greater than actual is greater than actual output for farm A, implying $\lambda_A = y_A / y_A^* < 1$. Since $n$ farms are present, a series of $n$ such linear programs must be solved to determine the technical efficiency of each farm.
Pure Technical Efficiency

The above problem imposes constant returns to scale on the technology by allowing the element of the intensity vector $z$ to take on any non-negative value. Afrait (1972) has shown that restricting the intensity vector to sum to one permits decreasing, constant and increasing return to scale. A new transformation set incorporating this non-constant returns technology can be expressed as:

$$T' = \{ (x, y) : x, y \leq Yz, z \in \mathbb{R}^k_+, \sum z_i = 1 \}$$

(4)

The new transformation set is illustrated in figure 2 for the original three farm A, B and C by the area under the curve $X_AABCT'$. Another efficiency measure, termed pure technical efficiency, PE can now be defined relative to this Frontier. For any particular observation $(x_i, y_i)$, pure technical efficiency can be expressed as:

$$PE(x_i, y_i) = \min \left( \lambda : (x_i, y_i, \lambda) \in T' \right)$$

(5)

Thus, PE equals one for all three observations in Figure 2, since all observations are on the technical frontier. To calculate this value numerically, the linear programming problem given by Eq. 3 is solved with the additional constraint that the elements of the intensity vector sum to one:

$$z_1 + z_2 + \ldots + z_n = 1$$

(3a)

Scale Efficiency

A measure of scale efficiency, SE, can now be defined as

$$SE(x_i, y_i) = \frac{TE(x_i, y_i)}{PE(x_i, y_i)}$$

(6)
If the technology exhibits constant return to scale at the observed input and output combination, the scale efficiency measure equals one. Value less than one indicate non-constant returns to scale. In Figure 2, farm B is the only scale efficient farm since both $\text{PE}_B(x_B, y_B)$ and $\text{TE}_B(x_B, y_B)$ are equal to one. While the pure technical efficiency measure $\text{PE}$ is equal to one for both farms A and C. The overall efficiency measure technical efficiency is less than one. The corresponding SE value thus be less than one, indicating the presence of non-constant returns to scale for these observation.

Farm A exhibits increasing return to scale, while farm C represents decreasing returns to scale. To determine the direction of non-constant returns to scale, is defined. This is accomplished by restricting the sum of the intensity variables to be less than or equal to one:

$$T^* = \{(x, y) : Xz \leq x, y \leq Yz, z \in \mathbb{R}^K, \Sigma z_i \leq 1\} \quad (7)$$

The non-increasing returns to scale technology frontier is illustrated in Figure 2 by the curve OBCT'. A new efficiency measure, WE*, referred to by Byrnes, Fare and Grosskopf (1984) as the weak star measure, can then be defined relative to the new transformation set $T^*$

![Figure 2: Transformation set (T) under non-constant returns to scale](image-url)
\[ W E^* (x_i, y_i) = \min (\lambda_i : (x_i, y_i, \lambda_i) \in T^*) \]  

The definition differs from the pure technical efficiency measure PE only in the inequality restriction on the summation constraint for the intensity variable element. Consequently, it is calculated by solving a third linear programming problem given by EQ. 3 and the following constraint, which replaces Eq. 3a;

\[
\sum z_i \geq 1 \\
\leq \\
= 
\]  

Since the restriction imposes non-increasing returns to scale, scale inefficiencies (SE 1) will be the result of decreasing returns to scale if only if \( WE^* > TE \). Such is the case for farm C. In contrast, farm A is experiencing returns to scale and thus \( WE^*_A = TE_A \).

**Congestion Efficiency**

Input congestion occurs when the marginal product of an input is negative. An example in agriculture is the application is the application of excessive fertilizer to field crops. To determine the effect of input congestion. The disposability states that output does not decrease for an increase in any input. Weak disposability differs in that output will not decrease for a proportional increase in all inputs. It thus permits the existence of cases in which an increase in a particular input will force production downward.

Under the assumption of weak disposability of inputs and non-constant returns to scale, a final transformation set \( T^{**} \) is defined as:

\[
T^{**} = \{(x, y) : Xz \leq Ox, y \leq Yz, z \in \mathbb{R}^k_+, \sum z_i = 1.0 < O \leq 1\} 
\]  

Where O permits the overutilization of inputs by relaxing the strong disposability assumption. Another measure of pure technical efficiency, $PE^*$, can now be derived relative to the frontier of this weakly disposable technology:

$$PE^* (x_j, y_i) = \min (\lambda i : (x_i, y_i, \lambda_i) \in T^{**})$$

(10)

This measure is calculated by solving the following linear programming problem:

$$\begin{align*}
\text{s.t.} & \\
\sum_{j=1}^{n} x_{ij} z_i & \geq \lambda x_i, \quad j = 1, 2, \ldots, n \\
\sum_{i=1}^{n} y_{ij} z_i - y_i \theta & \geq 0 \\
\sum_{i=1}^{n} z_i \theta & \\
0 & < \lambda \geq 1
\end{align*}$$

(11)

The effect of congestion or overutilization of a particular input on efficiency, $CE$, can then be determined by:

$$CE (x_i, y_i) = \frac{PE (x_i, y_i)}{PE^* (x_i, y_i)}$$

(12)

Congestion is evident for an individual farm if $CE$ is less than one. Overutilization of inputs is not present if the pure technical efficiency measures defined under weak ($PE^*$) and strong ($PE$) input disposability assumptions are equal.
Recent efforts to develop more general form of production function have focused almost exclusively on the elasticity of factor substitution. Several flexible functional forms includes the generalized Leontief (GL) the traslog (TL), the generalized Cobb-Douglas (GCD), the generalized square root quadratic (GSRQ), and the generalized Box-Cox (GBC). Since, by definition all these forms have the same attractive local property, it is not clear how the practitioner should choose among them.

Flexibility functional forms are attractive because they do not place any prior restrictions on consumer behavior at a base point. A precise statement of this flexibility property follows: For any particular data point (one observation on prices and income) any set of price and income elasticities can be achieved through an appropriate set of parameter value. A nonflexibility form is only capable of achieving a subset of the full range of price and income elasticities. A flexible form can achieve arbitrary price and income elasticities at any particular data point. However, once a set of parameter value is chosen, either a prior or statistical estimation, the pattern of price and income elasticities is determined for all possible data points.

Statistical analysis as the basis for choosing among functional forms. Implicit in theses statistical investigation is the recognition that ability of a function to perform well over range of a data points, in addition to the base point, is an important criterion to be used in choosing among flexible forms. This suggests that the global prosperities of flexible functional forms must be known before there can be adequate understanding of differences among these forms. To date, however, there has been no systematic theoretical analysis of the global properties of these forms. The existing evidence is based entirely upon the empirical results from specific data sets. Hence it is difficult to utilize the result to gain general insight into the properties of different flexible functional forms or to choose among forms for particular applications.
When selecting a functional form for use in empirical work, one is confronted by a choice between forms that exhibit good behavior globally and those that possess flexibility. Relatively simple forms, such as Cobb-Douglas and CES, satisfy certain regularity conditions globally, but the very simplicity that guarantees global good behavior also prevents such form modeling very sophisticated technologies. On the other hand, relatively complex forms having the flexibility to model fairly sophisticated technologies have been developed, but their very flexibility prevents them from being well-behaved globally. The current trend is clearly toward the development and use of flexible form which, although not globally well-behaved, may nonetheless satisfy the desired regularity conditions over a range of observations that contain or intersect the set of sample observations. In light of their inability to satisfy regularity conditions globally, and in light of the substantial econometric sophistication required in their estimation, it is worthwhile investigating just how well various flexible forms do model technology.

A difficulty with this empirical approach is that the true technology is unknown. Evaluating the performance of flexible forms on the basis of how well they fit observed data is useful if interest centers on the data, but may be misleading if interest centers on the functional forms themselves. An alternative approach, better suited to the evaluation of functional approach better suited to the evaluation of functional forms, is to begin with known technology and examine the ability of various forms to track that technology.

The Box-Cox transformation function to provide a variety of new generalized forms for direct and indirect production (utility) functions. These functional forms can be polynomials, quadratic in logarithms, or all sorts of mixed combinations of both. They also yield most of the commonly used flexible functional forms as special cases. The two-stage technology provide a general procedure for combining production functions or cost function to obtain new specifications suitable for empirical production analysis. Assumption of the production function analysis

Following are the important assumption of production function analysis:
1. The production function is defined only for non-negative values of inputs and outputs. In terms of (2.1), this means that
\[ y \geq 0 \]
\[ x_i \geq 0, \quad i = 1,2,\ldots,l \]

2. Every possible combination of inputs is assumed to result in maximum level of output. This means that the production function presupposes technical efficiency.

3. The input-output relationship or the production function is single valued and continuous for which there exist first and second order partial derivatives of the output, \( Y \), w.r.t. each of the input variables, \( X_1, X_2,\ldots,X_k \), i.e., \( \frac{\partial Y}{\partial x_i} \) and \( \frac{\partial^2 Y}{\partial x_i^2} \) (\( i = 1,2,\ldots,l \)) are non-vanishing.

4. The production function is characterized by (a) decreasing marginal product for all factor-product combinations, (b) decreasing rate of technical substitution between any two factors, and (c) an increasing rate of production transformation between any two products. Mathematically, this correspond to:

\[ \frac{\partial^2 y}{\partial x_i^2} < 0, \quad i = 1,2,3,\ldots,l \quad (2.13) \]

\[ \frac{\partial^2 x_i}{\partial x_j^2} < 0, \quad i, j = 1,2,3,\ldots,l \quad (i \neq j) \quad (2.14) \]

\[ \frac{\partial^3 y}{\partial y_i^2} > 0, \quad i, j = 1,2,3,\ldots,m \quad (i \neq j) \quad (2.15) \]

Interpreted differently, conditions (2.13)-(2.15) respectively require a concave production to the input axis, convex isoquants, and concave product transformation curves to the origin. Thus, we do not include cases of increasing and constant returns.
However, it is a common experience in agricultural production that as the level of input $X_i$ is increased, it results in increasing, constant and diminishing returns in sequence. Nevertheless, under usual conditions, the existence of diminishing returns to individual input variables is most likely.

5. The returns to scale are assumed to be decreasing. This means that a simultaneous one per cent increase in all the input variables results in less than one per cent increase in the output. In other words, doubling of all inputs less than doubles the output. Mathematically, this assumption can be stated as

$$\sum \left( \frac{x_i}{y} \right) \left( \frac{\partial y}{\partial x_i} \right) < 1, \quad i = 1,2,\ldots,, \ l \quad (2.16)$$

6. The exact nature of the firm's production function is assumed to be determined by a set of technical decision taken by the producer.

7. All the products and factors of production are perfectly divisible.

8. The parameters determining the firm’s production function do not change over the time period consideration. Besides, these parameters are not permitted to be random variables.

The method of production function analysis immediately breaks down if one or more of foregoing assumptions (1-8) do not hold.

Various methods of economic analysis have, however, been developed to meet situations when one or more of the foregoing assumptions are not meet. These alternative techniques of analysis often replacement of some of these assumptions by a new set of assumptions.
4.2.10.1 Linear functional form

Also known as first degree polynomial function, it is the simplest form of all the functions. Mathematically it is written as:

\[ Y = f(X; \beta) \]

Or

\[ Y = \beta_0 + \sum_{i=1}^{n} X_i \beta_i \]

Where;
- \( Y \) is the vector of output
- \( X \) is the matrix of input
- \( \beta \) is the vector of parameters

Characteristics:

The average product of \( X_1 \) input, can be obtained as

\[ AP_1 = \frac{X_1 \beta_1}{X_1} = \beta_1 \]

Thus, \( AP_1 \) remains constant irrespective of the level of input use.

The marginal product with respect to input \( X_1 \) is given by

\[ MP_1 = \frac{\partial X_1}{\partial Y} = \beta_1 \]

Thus for linear function \( AP = MP \) for all levels of input use.

The isoquant equation corresponding to the above production function for two input and representing a fixed level of output \( Y_0 \) is
$X_1 = f(X_2, Y_0)$

Here, the isoquant corresponding to a linear production function is also linear and downward sloping. It shows that $X_1$ and $X_2$ are perfect substitutes.

The elasticity of production $E_{pi}$ w.r.t. any input $X_i$ is unity, i.e.

$$E_{pi} = MP_i/ AP_i = \beta_i/\beta_i = 1$$

It indicates that one percent increase in the level of input $X_i$ causes exactly one percent increase in the level of output $Y$.

The rate of technical substitution of $X_2$ for $X_1$ input, is constant, i.e. $\beta_2/\beta_1$. The two inputs, therefore, are assumed to substitute at a constant rate whenever a linear production function is used.

**4.2.10.2 Quadratic Production Function**

Such functions are very popular with the researchers studying the effect of fertilizer/nutrient levels on the crop yields. The functional form is given by

$$Y = \beta_0 + \sum_{i=1}^{n} \beta_i X_i + \sum_{i=1}^{n} \beta_i X_i^2 + \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} X_i X_j$$

$i < j \land i, j = 1, 2, \ldots, n$

Where:

- $Y$ is the vector of output
- $X$ is the matrix of input
- $\sigma$ is the vector of parameters
- $\beta_0$ is the intercept coefficient
Characteristics:

A well behaved quadratic production function is concave to the input axis with convex isoquants and concave product transformation curves to the origin.

The average product w.r.t. input $X_1$ can be obtained as

$$\text{AP}_1 = Y/X_1 = \beta_1 + \beta_{11} X_1$$

The $\text{AP}_1$ corresponding to a quadratic production function with $\beta_{11} < 0$, which is monotonically declining linear curve. It can assume both positive and negative values.

The marginal product with respect to input $X_1$ is given by

$$\text{MP}_1 = \frac{\partial Y}{\partial X_1} = \beta_1 + 2\beta_{11} X_1$$

For the normal case where $\beta_{11} < 0$, $\text{MP}_1$ is also a monotonically declining linear curve like $\text{AP}_1$ with double rate of decline. $\text{MP}_1$ is both declining and negative, but can not be increasing and decreasing.

The isoquant equation can be derived from the quadratic production function with two input variables as follows:

$$X_1 = -(\beta_1 + \beta_{11} X_1) \pm \left[ (\beta_1 + \beta_{11} X_1)^2 - 4\beta_{11}(\beta_0 + \beta_2 X_2 + \beta_{22} X_2^2 - Y^0) \right]^{\frac{1}{2}}$$

The isoquants are not asymptotic to the input axes. Some isoquants intersects the input axes, and hence there corresponding levels of output can be achieved by using only one of the two factors, depending on the magnitude of $\beta_0$, $\beta_1$ and $\beta_{11}$ the isoquant corresponding to a distinct peak of the production function, denoting the maximum
output for a single combination of $X_1$ and $X_2$ inputs is represented by a point on the isoquant map called “Von Leibig Point”.

The rate of technical substitution of input $X_2$ for input $X_1$ can be obtained by taking the first order derivative of $X_1$ w.r.t. $X_2$ and then placing negative sign before it.

$$\frac{\beta_2 + 2\beta_{21} X_2 + \beta_{12} X_1}{\beta_1 + 2\beta_{11} X_1 + \beta_{12} X_2}$$

**Advantage:**

1. It is easy to restrict to exhibit constant or increasing marginal rate of transformation between outputs.

2. Quadratic function permits linear or convex transformation surfaces.

### 4.2.10.3 Cobb-Douglas Production functions

Cobb-Douglas or power function is a nonlinear production function. It is one of the most widely used functions in the economic analysis of the problems relating to empirical estimation in agriculture and industry. The functional form is given by

$$Y = \beta_0 \prod_{i=1}^{n} X_i^{\beta_i} \quad i = 1, 2, \ldots \ldots n$$

Where:

- $Y$ is the vector of output
- $X$ is the matrix of input
- $\beta$ is the vector of efficiency parameters and represent the production elasticities of the respective input variable
- $\beta_0$ is the intercept coefficient
Characteristics:

Average product w.r.t. $i^{th}$ input (AP$_i$) at given levels of all other inputs can be derived as

$$AP_i = \frac{Y}{X_i}$$

Thus it is seen that the average product of input $X_i$ is the function of its own level.

Marginal product w.r.t. $i^{th}$ input (MP$_i$) at given levels of all other inputs can be derived as

$$MP_i = \beta_1 X_i^{\beta_1} \beta_0 X_1^{\beta_1} X_2^{\beta_2} \ldots X_i^{\beta_1} X_{i+1}^{\beta_{i+1}} \ldots X_n^{\beta_n}$$

Which on solving gives

$$= \beta_1 \frac{Y}{X_i}$$

like $AP_i$, $MP_i$ is also a function of $X_i$. $MP_i$ must always be non-negative and decreasing. This further supports the flattening property of the production function which does not attain a finite maximum. Consequently, this function allows either constant, increasing or decreasing marginal productivity, and not all the three or even any two at the same time.

The isoquant equation $x_1 = g(x_2)$ is given as

$$x_1 = \left[ \frac{y_0}{(\beta_0 x_2^{\beta_2} x_3^{\beta_3} \ldots x_n^{\beta_n})^{1/\beta_0}} \right]$$

where $x_2, x_3, \ldots, x_n$ are held constant at specified levels. The resources serve as limitation inputs, i.e. the output become zero when any one of the inputs assumes a zero value. The isoquants are negatively sloped throughout and strictly convex to the origin, for positive values of inputs in the rational zone of production.
Rate of technical substitution (RTS) of $j^{th}$ input for $i^{th}$ input is given by

$$\text{RTS}_{ij} = \frac{\beta_j X_i}{\beta_i X_j}$$

RTS remains constant whenever $X_i$ and $X_j$ are increased in constant proportions, even though the output level increases.

Elasticity of production w.r.t. any variable input, say $X_i$, can be obtained as:

$$E_{pi} = \frac{\partial Y}{\partial X_i} = \beta_i$$

Thus, it is seen that the power of the respective inputs variables directly gives the elasticity of production w.r.t. it. The elasticity coefficients w.r.t. each of the input variables are, therefore, constant, irrespective of the input or output levels.

Elasticity of Substitution between any set of two input variables is constant and unity. This is the biggest limitation in using the above function.

Returns to Scale of this function is obtained as follows:

$$\text{Returns to Scale} = \sum (\frac{\partial y}{\partial x_i}) \left( \frac{x_i}{y} \right) = \beta_1 + \beta_2 + \ldots + \beta_n = \sum_{i} \beta_i$$

The summation of the powers of all the input variables provides us directly with a ready estimate of the returns to scale as also the degree of homogeneity of the production function. The return to scale are decreasing, constant or increasing, depending on whether $\beta_i$ is less than, equal to greater than one.
Advantages:

The main reason for the popularity of this function is not so much the conviction that the users of this function have about the realism and usefulness of the assumptions and conclusions that this function represents, but rather the ease with which this function can be estimated. This function provides a compromise between (a) Adequate fit of data, (b) computational feasibility, and (c) Sufficient degree of freedom unused to allow for statistical testing. Some of the other factors which favor the use of this function are:

1. It automatically:

   (a) Allow diminishing productivity of each of the resources which is a well recognized fact in agriculture; and

   (b) Causes the productivity of one resource to be dependent on the magnitude of others.

2. It does not require the estimation of too many coefficients for a single input factor and is economical in the use of degree of freedom in contrast to other functional forms.

3. Product contours under Cobb-Douglas equation become asymptotic to the axes and, therefore, may best apply to pairs of factors which serve as substitutes within one range but as technical complements for extreme ranges.

4. There is no prior specification regarding the returns to scale. The sum of regression coefficient in the unrestricted Cobb-Douglas function automatically indicates the returns to scale.

5. Coefficients in the Cobb-Douglas function are the production elasticities of inputs, which are independent of units of measurement and the factor ratios, and are easier in their manipulation and interpretation.
Disadvantage:

In spite of the above favorable points, there are some serious shortcomings of Cobb-Douglas production function as a tool for measuring the level of economic efficiency.

1. Cobb-Douglas function assumes constant partial and total elasticities of productions and Unit elasticity of substitution among factors of production. But actual conditions seldom conform to these basic assumptions of the model.

2. Recent studies indicate the existence of constant returns to scale. But as R. K. Sampath has shown, the assumption of perfect competition and profit maximization are inconsistent if the production function exhibits constant or increasing returns.

3. Single equation least square is the most popular method of estimating the parameters of Cobb-Douglas function. These estimated parameters suffer from three major econometric problems; viz. misspecification bias problem, simultaneous equation bias problems aggregation bias problem and the few other minor problems concerning the use of geometric mean. Another drawback in using geometric mean is that it divides the whole set of observations into two groups; one which are over allocating the resource concerned and the other which are under allocating it. This way it makes almost all the farmers inefficient.

4. Finally one of the important drawback of the Cobb-Douglas production function approach is that it attributes the prevailing level of economic inefficiency totally to the individual. This is inherent in the approach because it does not distinguish clearly the individual from the environment that he faces, within which he is supposed to take decisions which are to a great extent influenced by the characteristic of the environment.
5. This cannot be used satisfactorily for data where there are ranges of both increasing and decreasing marginal productivity. Neither can the function be used for data which might have both positive and negative marginal product.

6. The curve “flattens out” as input increases so a maximum product is not defined.

### 4.2.10.4 Constant Elasticity of Substitution Production Function (CES)

This function was first developed by Arrow, Chenery, Minhas and Solow (1961). It is the modification of Cobb-Douglas production function allowing elasticity of substitution to be more than unity, though constant. It is seen that the CES function is homogeneous of degree one. The two input functional form is given by

\[
Y = a \left[ \delta x_1^{-\rho} + (1-\delta)x_2^{-\rho} \right]^{1/\rho}
\]

Such that \( a > 0, \ 0 < \delta < 1 \) and \( \rho > -1 \)

Where \( x_1, x_2 \) are the quantities of input.
\( \delta, \rho \) and \( a \) are three parameters
also ‘\( a \)’ is called efficiency parameter; \( \delta \) is called the distribution parameter and \( \rho \) is called substitution parameter.

**Characteristics:**

The average product of \( X_1 \) input, can be obtained as

\[
AP_1 = x_1^{-1}y = x_1^{-1}a \left[ \delta x_1^{-\rho} + (1-\delta)x_2^{-\rho} \right]^{-1/\rho}
\]

The average product equation is homogeneous of degree one.
The marginal product with respect to input $x_1$ and $x_2$ can be found out by taking the first order derivatives of $y$ w.r.t. $x_1$ and $x_2$, respectively. The marginal product equation is given by:

$$MP_1 = \frac{\delta y}{\delta x_1} = \delta a \left[ (1-\delta)x_2^{-\rho} \right] \left( \frac{y}{x_1} \right)^{1+\rho}$$

these marginal products are homogeneous of degree zero. CES production function is characterized by diminishing returns to each of the inputs $x_1$ and $x_2$ for all positive levels of these inputs.

The isoquant equation for the production function is derived as

$$x_1 = \frac{(y/a)^\rho - (1-\delta)x_2^{-\rho}}{\delta}$$

the particular shape of the convex isoquant generated by a CES production function depends on the magnitude of elasticity of substitution. The isoquant of this function is always negatively sloped and convex to the origin in the rational zone of production. The elasticity of production is not a constant, but is the function of the levels of input with respect to which it is measured, keeping the other output level unchanged.

The Elasticity of production with respect to $x_1$ and $x_2$ inputs are

$$E_{p1} = \left( \frac{\delta y}{\delta x_1} \right) \frac{x_1}{y}$$

$$= \frac{1-\delta}{a^\rho} \left( \frac{y}{x_2} \right)^\rho$$

thus, elasticity of production is not constant, but is a function of the level of the input w.r.t. which it is being measured, keeping the other input level unchanged. Like marginal
product estimates, the elasticity of production estimates obtained from the CES production function also assume only positive values.

Elasticity of Substitution is constant for this function as the name implies but is not unity. When \( \rho = 0 \) this function reduces to the Cobb-Douglas function. The formula for elasticity of substitution (\( \sigma \)) is given by

\[
\sigma = \frac{[d \{ \ln (x_1/x_2) \}]}{[d \{ \ln (p_1/p_2) \}]}
\]

\[
\sigma = \frac{1}{1+\rho}
\]

is constant and whose its magnitude depends on the value of the parameter \( \rho \) as follows

\[
-1 < \rho < 0 \quad \rho = 0 \quad \Rightarrow \quad \begin{cases} 
\sigma > 1 \\
\sigma = 1 \\
\sigma < 1 
\end{cases}
\]

The general applicability of this function is quite restricted because of the nonlinear estimation problems and the necessity to select one among several alternative CES forms based on functional separability.

Limitations:

1. The main weakness of CES production is seems to be its empirical starting point. The derivation of the CES function is based on the assumption that the partial regression coefficient of input ratio is equal to zero. The assumption does not seems to be realistic. If this assumption does not hold, the value of the elasticity of substitution derived from the estimated CES function may be biased. Therefore
a general form of production function that does not depend upon this assumption is derives which includes CES function as a special case.

2. The function has restrictive priori-assumption on the parameter such as the elasticity if assumptions.

3. This function is non-linear in parameter and hence difficult to estimate.

4. It is difficult to generalize the function for more than two input case.

4.2.10.5 Variable Elasticity of Substitution Production Function (VES)

This function is generalization of Constant Elasticity of Substitution Production Function and the Cobb-Douglas production function. The function allows the elasticity of substitutions and returns to scale to vary with the level of output and input use and the same time exhibit the characteristics of homogenous class of production function.

There are two approaches of development of this production function

- Elasticity of factor substitution is function of capital intensity.
- Elasticity of factor substitution is a linear function of time.

The mathematical form is given by

\[ Y = a \left[ \delta x_1^\rho + (1-\delta)\eta \left( x_1 / x_2 \right)^{c(1+\rho)} \left( x_2^\rho \right) \right]^{-1/\rho} \]

Such that \( a > 0, \ 0 < \delta < 1 \) and \( \rho > -1 \)

Where \( x_1, x_2 \) are the quantities of input.
\( \delta, \rho \) and \( a \) are three parameters
also ‘a’ is called efficiency parameter, \( \delta \) is called the distribution parameter and \( \rho \) is called substitution parameter.
To estimate the parameters it is useful to write this function as:

\[ \ln Y = \ln a + \frac{1}{1+p} \ln x_1 + \frac{p}{1+p} \ln \left[ x_2 + \frac{8}{1+p} x_i \right] \]

**Characteristics:**

Marginal product equation of VES function with respect to input \( x_1 \) is given by

\[ MP_1 = \frac{\partial y}{\partial x_1} = \frac{1}{1+p} y + \frac{p}{1+p} \left( \frac{8}{1+p} x_i \right) \]

The function is homogeneous of degree one with positive marginal product.

In case of VES function every point on a given substitution curve has a different elasticity and hence called Variable Elasticity of Substitution Production Function. This assumption is the relaxation made in Constant Elasticity of Substitution Production Function. The elasticity varies with \( x, \left( \frac{x_1}{x_2} \right) \) ratio.

**Limitation:**

1. This function is non-linear in parameter and hence difficult to estimate.

2. It is difficult to generalize the function for more than two input case.

**4.2.10.6 Translog Production Function**

This is more flexible form of production function. This form of production function is becoming quite popular with the economist as it can be viewed as an exact representation of the production technology. Christensen, Jorgenson and Lau (1972) first developed this function. The translog function with \( n \)-inputs variables can be written as
Y = f(x_1, x_2, ..., x_n)

\[ Y = a_0 \prod_{i=1}^{n} x_i^{a_i} \prod_{i=1}^{n} x_i^{1/2 \sum_{j=1}^{n} b_{ij} \log x_i} \quad I, j = 1,2, ..., n \]

or the function in its log form can be written as

\[ \log y = \log a_0 + \sum_{i=1}^{n} a_i \log x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} \log x_i \log x_j \]

Where:

Y is the output
X is the input
a and b is the efficiency parameters
a_0 is the intercept coefficient

Characteristics:

Marginal product w.r.t. i^{th} input (MP_i) at given levels of all other inputs can be derived as

\[ MP_i = \left( \frac{\partial y}{\partial x_i} \right) \left( \frac{x_i}{y} \right) = \left[ a_i + \frac{1}{2} \sum_{j=1}^{n} b_{ij} \ln x_j \right] \left( \frac{y}{x_i} \right) \]

it is seen that for finite levels of X_i input, MP_i can be positive for a range of values of X_j, but can be negative if b_{ij} > 0 (for all i, j) and x_j tends to 0. similarly, if there exist at least one b_{ij} < 0, then MP_i < 0 as x_j tends to \infty. Thus, monotonicity requires that for all i, MP_i > 0, the translog function is not globally well behaved. It is a limitation of the translog production function.
Elasticity of production of the translog function w.r.t. input $X_i$ is obtained as

$$E_{pi} = \left( \frac{\partial \ln y}{\partial \ln x_i} \right)$$

$$= a_i + \frac{1}{2} \sum_{j=1}^{n} (b_{ij} \ln x_j)$$

Elasticity of production w.r.t. the $i^{th}$ input is not a constant but varies with the level of input $j$ ($j = 1, 2, \ldots, n$).

Return to scale are invariant with respect to the initial input levels and are equal to the sum of the production elasticities. Thus, for a translog function, which is non-homogeneous, the returns to scale (function coefficient) are not invariant with input levels. The functional coefficient are given by

$$\xi = \sum_{i=1}^{n} E_{pi} = \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} \sum_{j=1}^{n} (b_{ij} \ln x_j)$$

The function allows the researcher to estimate the nature of return to scale and factor substitution instead of imposing the priori restriction on them. It exhibits non-constant elasticity i.e., increasing, decreasing and negative marginal returns singularly, in pairs or all three simultaneously. Thus all three stages of production are visible here.

Advantages:

1. Although it can be viewed as an exact representation of the productive technology the translog function is best seen as a second-order Taylor series expansion around the mean of output and each input.
2. It is chosen because it allows the researcher to estimate the nature of returns to scale and factor substitution, instead of imposing a priori restriction.

3. Furthermore, regions with non-convex isoquants (non-economic region) exist whenever any coefficient is estimated to be significantly different from zero. (If non is the function reduces to the Cobb-Douglas form.

4. Neutral technical progress is easily taken into account. The traslog function is not locally constraint by assumptions of homogeneity or additiviy.

5. The function exhibit non constant elasticity that is increasing, decreasing and negative marginal returns singularly, in pairs, or all three simultaneously.

**Disadvantages** :

1. As the number of variable increases, the number of parameters to be estimated increase in large number for eg. with three variable case there are ten parameter to be estimated hence with the increase in number of variables the estimation of parameters becomes very cumbersome.

2. The problem of multicollinearity is serious with traslog function.

### 4.2.10.7 Generalized Leontief Production Function

This is a flexible functional form placing no priori restriction on Allen partial elasticity of substitution. Diewert (1971) developed this. This function is linear in parameters hence easy to compute. The functional form is given by

\[
Y = \sum \sum a_{ij} x_i^{1/2} x_j^{1/2} + \sum a_i x_i
\]
for all values of $a_{ij} = a_{ji}, a_{ij} \geq 0$

$i, j = 1, 2, \ldots, n$ and $i \neq j$

**Characteristics:**

The function is a concave function.

Generalized Leontief production is a non-decreasing function in $x$ follows from the non-negativity of the parameter $a_{ij}$.

If $a_{ij} = 0$ for $i \neq j$, then the function reduces to a linear production function.

The approximation is non-homothetic unless $a_i = b_i = 0 \ \forall i$, in which case it is linearly homogeneous.

The return to scale equation is derived as:

$$
\left( \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i^{1/2} x_j^{1/2} + 2 \sum_{i=1}^{n} a_i x_i \right)^{-1}
$$

The partial elasticities of substitution are given by

$$
\sigma_{ij} = (\frac{1}{2} a_{ij} x_i^{-1/2} x_j^{-1/2}) + 2 \sum_{i} a_i x_i + \sum_{i} b_i x_i
$$

The form can achieve arbitrary elasticities at any particular data point.

The production function is quasi-concave in nature following all the assumptions of neoclassical production theory.
It can also be derived as a limiting case of Generalized Box-Cox model. However the Generalized Box-Cox has the disadvantage of not being a true maximum likelihood estimator since the distribution of the disturbance is not normal as it is assumed to be.

Limitations:

The Generalized Leontief Production Function approximation is incapable of distinguishing among homotheticity, homogeneity, and linear homogeneity. It collapses to a fixed proportions forms if $a_i = 0 \forall i \neq j$.

4.3 EMPIRICAL ANALYSIS

Considering the above theoretical background of estimation and specification of technology frontier model, the frontier production functions of different functional forms were estimated considering the following specification;

\[ Y = f( L, N, K) \]

Where;

\[ Y = \text{is the vector of output on the farm in kg} \]
\[ L = \text{is the vector of land in hectare} \]
\[ N = \text{is the vector of labor in man days} \]
\[ K = \text{is the vector of capital in Rupees} \]

The data has been classified into two periods 1985-90 and 1990-95 representing the technology-1 and technology-2, respectively. The production function of the above specifications were estimated for working out efficiency. The indices of technical efficiency were computed by using the following formula:
\[ Y_i = f(X_i, \beta) + \varepsilon_i \]

Then

\[ TE_i = Y_i / (f(X_i, \beta) + \varepsilon_i) \]

such that \( 0 \leq TE_i \leq 1 \)