# Monte Carlo Simulation Applied to Stochastic Frontier Production Function in the Parameter Estimation

V. S. Mathu Suresh[1], Dr. G. Easwara Prasad[2]

[1]Assistant Professor, Department of Mathematics, Udaya School of Engineering, Ammandivillai, Kanyakumari District, INDIA

[2]Professor and Head, Department of Mathematics, S. T. Hindu College, Nagercoil, Kanyakumari District, INDIA

***Abstract*– Several studies have expressed different forms of production functions and frontier production functions with different distribution structures of each one of the error term. The maximum likelihood parameter estimation gives problems in the estimation due to the non linearity. Here an attempt is made to use the simulation exercise in the estimation of parameters. As the sample size and the replications are increased the estimates converge to the real value.**

***Keywords*– average production, random error, estimation, relative variability.**

## I. Introduction

The theoretical definition of a production function expressing the **maximum** amount of output obtainable from a given input bundles with fixed technology has been accepted for many decades. And for almost as long, the researchers have been using the average production functions. The concept of frontier production function is introduced to bridge the gap between the theory and the empirical work. The latest in this is the splitting of the error term into two one is set to be the normal and the other has one – sided distribution. This has helped to overcome the shortcoming of the previous research work.

## II. Parametric frontier models

For a given firm ($i^{th}$ firm) we write

$$y_i = f(x_i\,;\,\boldsymbol{\beta}) \qquad (1)$$

as the production relation. In this $y_i$ is the maximum output obtainable from $x_i$, a vector of (non- stochastic) inputs, and β is the unknown parameter vector to be estimated.

The mathematical programming approach of estimation of β is based on a cross section of N firms within a given industry. It is by minimizing

$$\sum |y_i - f(\boldsymbol{x}_i\,;\,\boldsymbol{\beta})|\,,$$

Subject to $y_i \le f(\boldsymbol{x}_i\,;\,\boldsymbol{\beta})$ ,

Which is a LP if $f(\boldsymbol{x}_i\,;\,\boldsymbol{\beta})$ is linear in **β**.

Alternatively if instead of mod we use the square subject to the same constraint, it is quadratic programming provided $f(\boldsymbol{x}_i\,;\,\boldsymbol{\beta})$ is linear.

One major drawback in these is that these are highly sensitive to outliers. This has made the researchers to go for the " Probabilistic" frontiers. Here the estimation is the same as the mathematical programming (LP) techniques discussed above, the only addition is that a **specified** proportion of observations is allowed to lie above the frontier. The selection of this proportion is essentially arbitrary, lacking explicit economic or statistical justification. Another problem involves reconciling the observations above the frontier with the concept of frontier with maximum possible output. This is accomplished by appealing to measurement error in extreme observations. Still, it seems preferable to incorporate the possibility of measurement error, and of other unobservable Shocks, in a less arbitrary fashion.

Since the mathematical programming results do not satisfy the statistical properties, it is decided to add a one – sided disturbance to (1) with which the model becomes

$$y_i = f(\boldsymbol{x}_i\,;\,\boldsymbol{\beta}) + \epsilon_i\,;\; i = 1,2,\dots,N \qquad (2)$$

Where $\epsilon_i \le 0$ . After prescribing the distribution assumption for $\epsilon_i$, the model can be estimated by the maximum – likelihood techniques. In the particular case when $-\epsilon_i$ has an exponential distribution it leads to the LP techniques, on the other hand $-\epsilon_i$ is half – normal leads to Quadratic Programming (QP) technique. The estimates through the above methods discussed since $y_i \le f(\boldsymbol{x}_i\,;\,\boldsymbol{\beta})$, the random variable y depends on the parameters to be estimated and hence the asymptotic properties of the behavior of the distribution of the parameters cannot be done. Hence the full knowledge of the frontier after estimation cannot be done.

For this purpose the same set of Research group constructed an alternate error structure as follows:

$$\epsilon_i = \begin{cases} \epsilon_i^* / \sqrt{1-\theta}, & if \quad \epsilon_i^* > 0, \quad i = 1,2,\dots,N \\ \epsilon_i^* / \sqrt{\theta}, & if \quad \epsilon_i^* \le 0, \end{cases} \qquad (3)$$

Where $\epsilon_i^*$ are independent normally distributed random variables with zero mean and constant variance $\sigma^2$.

For $0 < \theta < 1$ ; otherwise , $\epsilon_i^*$ has either the negative or positive truncated normal distribution, when $\theta = 1$ or $\theta = 0$ , respectively.

Here the justification is that the firms are pressured to differ in their 'production' of $y$ for a given set of values for the 'inputs' according to random variation in (1) their ability to utilize 'best practice' technology, a source of error that is one – sided ($\epsilon_i \le 0$), and/or (2) an input quantity or measurement error in $y$ , a symmetric error. The parameter $\theta$ is interpreted as the measure of 'relative variability' in those two error sources. its values circumscribing the full frontier function ($\theta = 1$), the 'average' function $\left(\theta = \frac{1}{2}\right)$ , and intermediate cases of some interest.

In the above if $\theta \to 1$ , the positive error component has a large variance and hence it has small influence in the likelihood function and the negative error dominates. This gives rise to the ''full'' frontier as the limiting case ($\theta = 1$). A similar interpretation follows for $\theta \to 0$ . When $\theta = \frac{1}{2}$ , the likelihood has the form of a mixture of two half-normal's , each with equal influence.

## III. A STOCHASTIC FRONTIER

Here the basic model is (2)

$$y_i = f(\boldsymbol{x_i} ; \boldsymbol{\beta}) + \epsilon_i , i = 1,2,\dots,N \qquad (4)$$

Where $\epsilon_i = u_i + v_i$

Here $v_i$ is the symmetric disturbance : ie, $\{v_i\}$ are assumed to be $i.i.d. N(0, \sigma_v^2)$ . $u_i$ is assumed to be distributed independently of $v_i$, and to satisfy $u_i \le 0$. Here we are particular to the case wherein $u_i$ is assumed to be $N(0, \sigma^2)$ which is truncated about at zero. One other which is also tenable is the exponential distribution for $-u_i$ . When $\sigma_v^2 = 0$ , the model becomes a deterministic frontier model , $\sigma_u^2 = 0$ make it the usual stochastic production function model.

The non-positive distribution $u_i$ reflects the fact that each firm's output must lie on or below its frontier $[f(\boldsymbol{x_i} ; \boldsymbol{\beta}) + v_i]$. Any such deviation is the result of factors under the firms control , such as technical and economic inefficiency , the will and effort of the producer and his employee, and perhaps such factors as defective and damaged product. But the frontier itself can vary randomly across firms, or over time for the same firm. On this interpretation, the frontier is stochastic, with random distribution $v_i >/< 0$ being the result of favourable as well as unfavourable external events such as luck, climate topography, and machine performance. Errors of observation and measurement on $y$ constitute another source of $v_i >/< 0$.

One additional advantage in this approach is that we can estimate the variances of $v_i$ and $u_i$ , so as to get evidence on their relative sizes. Another implication of this approach is that productive efficiency should , in principle, be measured by the ratio

$$y_i \; / \; [f(\boldsymbol{x_i} ; \boldsymbol{\beta}) + v_i] \qquad (5)$$

rather than by the ratio

$$y_i \; / \; [f(\boldsymbol{x_i} ; \boldsymbol{\beta})] , \qquad (6)$$

This simply distinguishes productive inefficiency from other sources of disturbance that are beyond the control of the firm's. For example, a farmer whose crop is decimated by drought or storm is unlucky on our measure (5) , but inefficient by the used measure (6).

To simplify the discussion on the estimates, we consider the linear model. Now the equation in the matrix form is

$$\boldsymbol{y} = X\boldsymbol{\beta} + \epsilon , \qquad (7)$$

Instead of (2) , now $\epsilon = \boldsymbol{v} + \boldsymbol{u}.$

## IV. ESTIMATION OF THE STOCHASTIC FRONTIER MODEL

The density function of $\epsilon$ is

$$f(\epsilon) = \int_0^\infty f(u, \epsilon) \, du$$

Where

$$f(u, \epsilon) = \frac{1}{\pi \, \sigma_u \sigma_v} \exp\left\{\frac{-u^2}{2\,\sigma_u^2} - \frac{(\epsilon + u)^2}{2\,\sigma_v^2}\right\}$$

on integration w.r.to u it gives

$$f(\epsilon) = \frac{2}{\sigma} f^*\left(\frac{\epsilon}{\sigma}\right) [1 - F^*(\epsilon \lambda \sigma^{-1})] , -\infty \le \epsilon \le \infty \qquad (8)$$

Where $\sigma^2 = \sigma_u^2 + \sigma_v^2$. $\lambda = \sigma_u / \sigma_v$ and $f^*(\cdot)$ and $F^*(\cdot)$ are the standard normal density and distribution functions respectively. This density is asymmetric around zero and it's mean and variance are

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

$$V(\epsilon) = V(u) + V(v)$$

$$= \left(\frac{\pi - 2}{\pi}\right) \sigma_u^2 + \sigma_v^2 \qquad (9)$$

Here $\lambda$ is an indication of the relative variability of the two sources of random error that distinguish firms from one another. $\lambda^2 \to 0$ implies $\sigma_v^2 \to \infty$ and / or $\sigma_u^2 \to 0$ , that is the symmetric error dominates in the determination of $\epsilon$. In this case (8) becomes the density of a $N(0, \sigma^2)$ random variable. Similarly when $\sigma_v^2 \to 0$, the one sided error becomes the dominated source of random variation and (5) is then the negative half-normal.

The estimation is done by maximizing the likelihood.

$$\ln L\left({}^{y}/_{\boldsymbol{\beta}}, \lambda, \sigma^2\right) = N \ln\frac{\sqrt{2}}{\sqrt{\pi}} + N \ln \sigma^{-1} + \sum_{i=1}^{N} \ln[1 - F^{*}(\epsilon_i\, \lambda\, \sigma^{-1})] - \frac{1}{2\sigma^2}\sum_{i=1}^{N} \epsilon_i^2 \qquad (10)$$

With respect to $\sigma^2$ , $\lambda$ and $\boldsymbol{\beta}$ . Here $N$ is the size of the sample. Various solution algorithms are available for solution.

## V. Monte Carlo Application

To find some specific information about small sample behavior of the ML estimation already discussed, we constructed two limited Monte Carlo experiments fully by using simulation data for each variable within specified limits.

### A. Illustration: I

The model considered is $y_i = \epsilon_i$ $(i = 1,2, \dots, N)$

Monte Carlo results for the model $y_i = \epsilon_i$ ; number of replication 200 sample size 100 value of $\lambda =$ 1.845 , 1.342 , 0.838

TABLE 1

| 1.845 | 1.342 | 0.838 |
|---|---|---|
| $\hat\sigma^2 = 0.058$ (0.06)<br>(0.063) (0.23)<br>$\hat\lambda = 1.85$ $\hat\sigma_u^2 = 0.002$ (0.02)<br>(0.003) $\hat\sigma_v^2 = 0.001$ (0.04)<br>(0.001) | $\hat\sigma^2 = 0.058$ (0.05)<br>(0.057) (0.021)<br>$\hat\lambda = 1.36$ $\hat\sigma_u^2 = 0.0025$ (0.023)<br>(0.002) $\hat\sigma_v^2 = 0.001$ (0.038)<br>(0.0015) | $\hat\sigma^2 = 0.048$ (0.04)<br>(0.051) (0.018)<br>$\hat\lambda = 0.9$ $\hat\sigma_u^2 = 0.001$ (0.020)<br>(0.001) $\hat\sigma_v^2 = 0.002$ (0.035)<br>(0.002) |
| (0.003) $\hat\sigma^2 = 0.071$ (0.09)<br>(0.001)<br>(0.087) (0.21)<br>$\hat\lambda = 1.85$ $\hat\sigma_u^2 = 0.002$ (0.026)<br>(0.008) $\hat\sigma_v^2 = 0.001$ (0.044)<br>(0.005) | (0.002) $\hat\sigma^2 = 0.063$ (0.058)<br>(0.0015)<br>(0.079) (0.021)<br>$\hat\lambda = 1.36$ $\hat\sigma_u^2 = 0.002$ (0.029)<br>(0.006) $\hat\sigma_v^2 = 0.002$ (0.049)<br>(0.004) | (0.001) $\hat\sigma^2 = 0.051$ (0.048)<br>(0.002)<br>(0.073) (0.023)<br>$\hat\lambda = 0.8$ $\hat\sigma_u^2 = 0.002$ (0.026)<br>(0.002) $\hat\sigma_v^2 = 0.003$ (0.041)<br>(0.003) |

Same as above with same number of replication and sample size is increased to 200 with the same model.

TABLE 2

| 1.845 | 1.342 | 0.838 |
|---|---|---|
| $\hat\sigma^2 = 0.78$ (0.14)<br>(0.087) (0.20)<br>$\hat\lambda = 1.86$ $\hat\sigma_u^2 = 0.007$ (0.021)<br>(0.008) $\hat\sigma_v^2 = 0.004$ (0.045)<br>(0.005) | $\hat\sigma^2 = 0.070$ (0.06)<br>(0.079) (0.025)<br>$\hat\lambda = 1.39$ $\hat\sigma_u^2 = 0.004$ (0.03)<br>(0.006) $\hat\sigma_v^2 = 0.003$ (0.051)<br>(0.004) | $\hat\sigma^2 = 0.059$ (0.05)<br>(0.073)<br>$\hat\lambda = 0.8$ $\hat\sigma_u^2 = 0.001$ (0.02)<br>(0.002) $\hat\sigma_v^2 = 0.002$ (0.04)<br>(0.003) |

Where $\epsilon_i$ is generated by (8) with values for $\sigma_u^2$ and $\sigma_v^2$ in two different farm data already studied by two researchers (unpublished Ph.D thesis ).

The results and the details of the values of all the parameters used in the simulation are reported Table 1. The results indicate only very small bias and especially $\sigma_u^2$ , $\sigma_v^2$ are very well since MSE is very low in almost all the cases.

In Table 2 since sample size is doubled , the precisions are still higher for all , which indicates the importance of very large samples in simulation studies. One important thing is that no regression is needed.

---

- Values in parenthesis to the right side of the parameters are the MSE's. Values in parameters to the left of $\hat\sigma^2$ , $\hat\sigma_u^2$ , $\hat\sigma_v^2$ are the true values of the parameters used in simulation for all the 4 tables.

TABLE 3

| 1.845 | 1.342 | 0.838 |
|---|---|---|
| $\hat\sigma^2 = 0.054$ (0.08)<br>(0.063) (0.87)<br>$\hat\lambda = 1.84$ (0.11)<br>$\hat\mu = 0.8$<br>(0.003) $\hat\sigma_u^2 = 0.006$ (0.018)<br>(0.001) $\hat\sigma_v^2 = 0.005$ (0.050) | $\hat\sigma^2 = 0.051$ (0.08)<br>(0.057) (0.86)<br>$\hat\lambda = 1.34$ (0.17)<br>$\hat\mu = 0.83$<br>(0.002) $\hat\sigma_u^2 = 0.005$ (0.026)<br>(0.0015) $\hat\sigma_v^2 = 0.006$ (0.056) | $\hat\sigma^2 = 0.040$ (0.07)<br>(0.051) (0.84)<br>$\hat\lambda = 1.91$ (0.20)<br>$\hat\mu = 0.7$<br>(0.001) $\hat\sigma_u^2 = 0.003$ (0.013)<br>(0.002) $\hat\sigma_v^2 = 0.009$ (0.006) |
| (0.003) $\hat\sigma^2 = 0.061$ (0.013)<br>(0.001)<br>(0.087) (0.93)<br>$\hat\lambda = 1.91$ (0.18)<br>$\hat\mu = 0.8$<br>(0.008) $\hat\sigma_u^2 = 0.001$ (0.015)<br>(0.005) $\hat\sigma_v^2 = 0.001$ (0.032) | (0.002) $\hat\sigma^2 = 0.058$ (0.072)<br>(0.0015)<br>(0.079) (0.089)<br>$\hat\lambda = 1.57$ (0.23)<br>$\hat\mu = 0.8$<br>(0.006) $\hat\sigma_u^2 = 0.002$ (0.014)<br>(0.004) $\hat\sigma_v^2 = 0.002$ (0.031) | (0.001) $\hat\sigma^2 = 0.051$ (0.030)<br>(0.002)<br>(0.073) (0.06)<br>$\hat\lambda = 0.84$ (0.23)<br>$\hat\mu = 0.7$<br>(0.002) $\hat\sigma_u^2 = 0.004$ (0.011)<br>(0.003) $\hat\sigma_v^2 = 0.008$ (0.004) |

*Illustration : II*

Here the experiment is repeated by changing the model to $y_i = \mu + \epsilon_i$ $(i = 1,2,\dots,N)$. Already $E(\epsilon) = 0$ now the effect of $\mu$, the intercept is questionable and the results presented in Table 3 and Table 4 reveal the following:

Monte Carlo results for the model $y_i = \mu + \epsilon_i$; number of replication = 200, sample size is 100. $\mu = 1.0$, $\lambda = 1.845, 1.342, 0.838$*

Same as above with sample number and of replication and sample size is increased to 200 with the same model.

TABLE 4

| 1.845 | 1.342 | 0.838 |
|---|---|---|
| $\hat{\gamma}^* = 0.068$ (0.047)<br>(0.087) (1.11)<br>$\hat{\lambda} = 2.008$ (0.017)<br>$\hat{\mu} = 0.98$<br>(0.008) $\hat{\sigma}_u^2 = 0.004$<br>(0.005) $\hat{\sigma}_v^2 = 0.008$ | $\hat{\gamma}^* = 0.043$<br>(0.079) (0.82)<br>$\hat{\lambda} = 1.533$ (0.019)<br>$\hat{\mu} = 0.93$<br>(0.006) $\hat{\sigma}_u^2 = 0.002$<br>(0.004) $\hat{\sigma}_v^2 = 0.007$ | $\hat{\gamma}^* = 0.048$<br>(0.073) (0.36)<br>$\hat{\lambda} = 0.872$ (0.039)<br>$\hat{\mu} = 0.6$<br>(0.002) $\hat{\sigma}_u^2 = 0.001$<br>(0.003) $\hat{\sigma}_v^2 = 0.003$ |

In Table 3 the results for $\hat{\lambda}$ are altered somewhat with some cases of negative bias. Moreover , the additional parameter has reduced the values of $\hat{\sigma}_u^2$ and hence $\hat{\sigma}^2$. In the previous cases the estimates were very sharp and this change is due to the presence of $\hat{\mu}$. Another important achievement is that when the sample size is increased there is not much of change in the results in the error is also minimized. Now the result are comparable with that of Aigner Amemiya and Poirier (1976).

## VI. CONCLUSION

As illustrated the use of simulation in the parameter estimation from the ML for the stochastic frontier production function is a possibility with reduced calculation.

## REFERENCES


[1] Afriat, S.N. , 1972, Efficiency estimation of production functions, International Economic Review 13, Oct , 568 – 598.

[2] Aigner , D.J. and S.F. Chu , 1968, On estimating the industry production function, American Economic Review 58 , 826 – 839.

[3] Aigner, D.J. , T. Amemiya and D.J. Poinier , 1976 , On the estimation of Production frontiers , International Economic Review. 17, 377 – 396.

[4] Farell, M.J. , 1957, The Measurement Of Production efficiency, Journal of Royal statistical Society (A, general) 120, pt 3, 253 – 281.

[5] Hildebrand G. and T.C. Liu , 1965 , Manufacturing Production Functions in the United States , 1957 (Cornell University Press , Ithaca, N.Y ).

[6] Meera Krishnamoorthy.(1996),"an Econometric Analysis of measuring the technical efficiency in Agriculture in Tamilnadu". Un published Ph.d work Tamilnadu Agricultural University, Coimbatore.

[7] Schmidt.P., 1976, On The Statistical Estimation of Parametric frontier Production Functions, Review of Economics and Statistics 58, 238 – 239.

[8] Timmer, C. Peter, 1971, Using Probabilistic Frontier Production Function to Measure the Technical Efficiency, Journal of Political Economic 79, 776 – 794.

[9] Veerabathran.R.(2003),"Stochastic Process- An Application to Agriculture" un published Ph.D work Manmoniam Sundaranar University Tirunelveli.

[10] Zellner, A.Kmenta, J. and J.Dreze, 1966, Specification and estimation of Cobb-Douglas, Production Functions, Econometric a 24, Oct , 784 – 795.