5-2012

A Normal Truncated Skewed-Laplace Model in Stochastic Frontier Analysis

Junyi Wang
Western Kentucky University, junyi.wang905@topper.wku.edu

A NORMAL TRUNCATED SKEWED-LAPLACE MODEL IN STOCHASTIC FRONTIER ANALYSIS

A Thesis
Presented to
The Faculty of the Department of Mathematics and Computer Science
Western Kentucky University
Bowling Green, Kentucky

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Junyi Wang

May 2012

A NORMAL TRUNCATED SKEWED-LAPLACE MODEL IN STOCHASTIC FRONTIER ANALYSIS

Date Recommended 04/24/2012

Ngoc Nguyen

Dr. Ngoc Nguyen, Director of Thesis

Melanie A.

Dr. Melanie Autin

by direction of Tom Richmond N.N

Tom Richmond

Dr. Tom Richmond

Kinchel C Doerner 19-June-2012

Dean, Graduate Studies and Research Date

ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere appreciation and gratitude to my adviser, Dr. Ngoc Nguyen, for her continuous encouragement, guidance and support, without which the completion of this thesis would not have been possible.

I would also like to thank Dr. Melanie Autin and Dr. Tom Richmond, who have kindly agreed to be on my committee and to help me with their expertise on research design.

I am also deeply grateful to Dr. Claus Ernst, who has offered his generous assistance and advice during my completion of this master's degree program, and who has chose to stand firmly behind me when my confidence was weakened and my judgement was clouded.

Special thanks go to my dear parents and grandparents, who have loved me unconditionally and have always believed in me. I am also obliged to my cousin, with whom I have shared wonderful childhood memories. She has always been there for me during difficult times and has encouraged me to hold on to my dreams.

LIST OF FIGURES

CONTENTS

A NORMAL TRUNCATED SKEWED-LAPLACE MODEL IN STOCHASTIC FRONTIER ANALYSIS

Junyi Wang May 2012 40 Pages

Directed by: Dr. Ngoc Nguyen, Dr. Melanie Autin, and Dr. Tom Richmond

Department of Mathematics and Computer Science Western Kentucky University

Stochastic frontier analysis is an exciting method of economic production modeling that is relevant to hospitals, stock markets, manufacturing factories, and services. In this paper, we create a new model using the normal distribution and truncated skew-Laplace distribution, namely the normal-truncated skew-Laplace model. This is a generalized model of the normal-exponential case. Furthermore, we compute the true technical efficiency and estimated technical efficiency of the normal-truncated skewed-Laplace model. Also, we compare the technical efficiencies of normal-truncated skewed-Laplace model and normal-exponential model.

Chapter 1

Introduction

Stochastic frontier analysis (SFA) is a method of economic modeling, and it is widely used to estimate individual efficiency scores. The stochastic frontier analysis was first proposed by Aigner et al.(1977) and Meeusen and van de Broeck (1977). SFA modeling has been used in various fields, such as hospitals, stock markets, manufacturing factories, services, etc. The purpose of stochastic frontier analysis is to measure how efficient a producer is with given observations of input and output by using two error terms, u and v. This method is often found to be useful in estimating the values of function in profit and cost.

Several decades ago, many economists made an effort to analyze and compute the production efficiency by using the models of production, cost, and profit. They began with the model of production function which producers use to maximize the outputs obtainable from a given amount of inputs. In order to achieve the maximum production output level, one needs to minimize the cost function and maximize the profit function.

Cobb and Douglas (1928), Arrow et al. (1961), Berndt and Christensen (1973), and their followers have made certain new developments to the original production function; they consider the production function to be more flexible in that it needs to take into account some random noise that can affect the production process in reality. The random noise may come from weather changes, unpredictable variations in machines, economic adversities or labor performance, and so on. Thus the original production function is no longer appropriate because it is performed primarily based on the ideal conditions. As a result, econometricians come to define the new function as the production frontier.

Production frontier is defined as characterizing "minimum input bundles required to produce various outputs, or the maximum output producible with various input bundles, and a given technology"(Kumbhakar and Lovell, 2000). In other words, the production frontier refers to the minimum input needed for any given output, or the maximum output from any given input. Producers operating on their production frontier are called technically efficient, and producers operating beneath their production frontier are called technically inefficient.

There are two differences between the production function and production frontier. First of all, the production function is an ideal function without considering some random noise, while the production frontier does consider various practical applications. Secondly, the production function is symmetrically distributed with zero mean, but the composed error terms of production frontier must be a skewed variable with nonzero mean.

In summary, stochastic frontier analysis is a method used to analyze and compute technical efficiency based on the stochastic production frontier. The production function is an ideal model which does not consider the random noise, while the production frontier model includes various random noises reflecting more flexibility in real life.

1.1. Stochastic Production Function

The estimation of the parametric frontier production function is based on the pioneering work of Aigner and Chu (1986), Afriat (1972), and Richmond (1974). In a given industry, firms might have different production processes due to certain technical parameters in the industry, different scales of operation or organizational structures. In order to understand the problem, some fundamental definitions are presented below. First, we assume that producers use a nonnegative vector of inputs, denoted

$x = (x_1, x_2, ..., x_N) \in \mathbb{R}^N_+$, to produce a nonnegative vector of outputs, denoted $y = (y_1, y_2, ..., y_M) \in \mathbb{R}^M_+$.

Definition 1.1.1. *The graph of production technology*

$$GR = \{(y, x) : x \text{ can produce } y\}$$

describes the set of feasible input-output vectors.

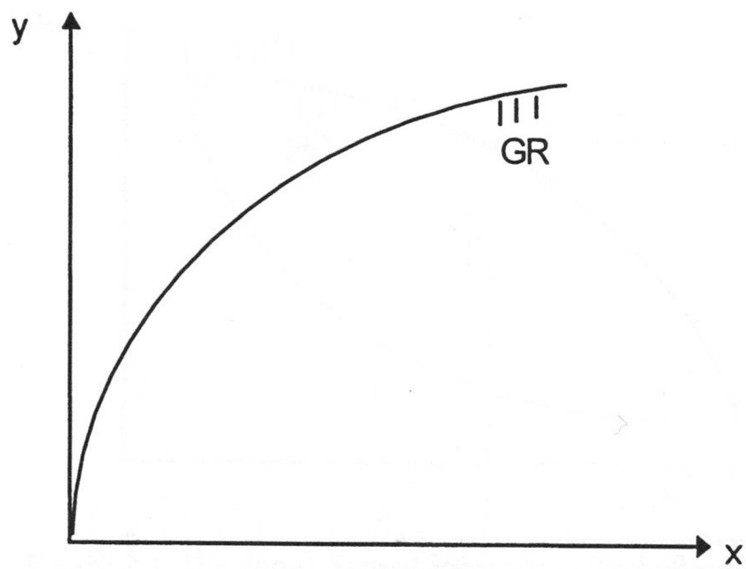

Figure 1.1.1. The Graph of Production Technology (M=1,N=1)

Figure 1.1.1 illustrates the graph of production with a single nonnegative input and a single nonnegative output. The graph is known as the production possibilities set.

Definition 1.1.2. *The output sets of production technology*

$$P(x) = \{y : (y, x) \in GR\}$$

describe the sets of output vectors that are feasible for each input vector $x \in \mathbb{R}^N_+$

Definition 1.1.3. *The output isoquants*

$$Isoq\ P(x) = \{y : y \in P(x), \lambda y \notin P(x), \lambda > 1\}$$

describe the sets of all output vectors that can be produced with each input vector x but which, when radially expanded, cannot be produced with input vector x.

DEFINITION 1.1.4. *The output efficiency subsets*

$$Eff\ P(x) = \{y : y \in P(x), y' \geq y \Rightarrow y' \notin P(x)\}$$

describe the sets of all output vector that can be produced with each input vector x but which, when expanded in any dimension, cannot be produced with input vector x.

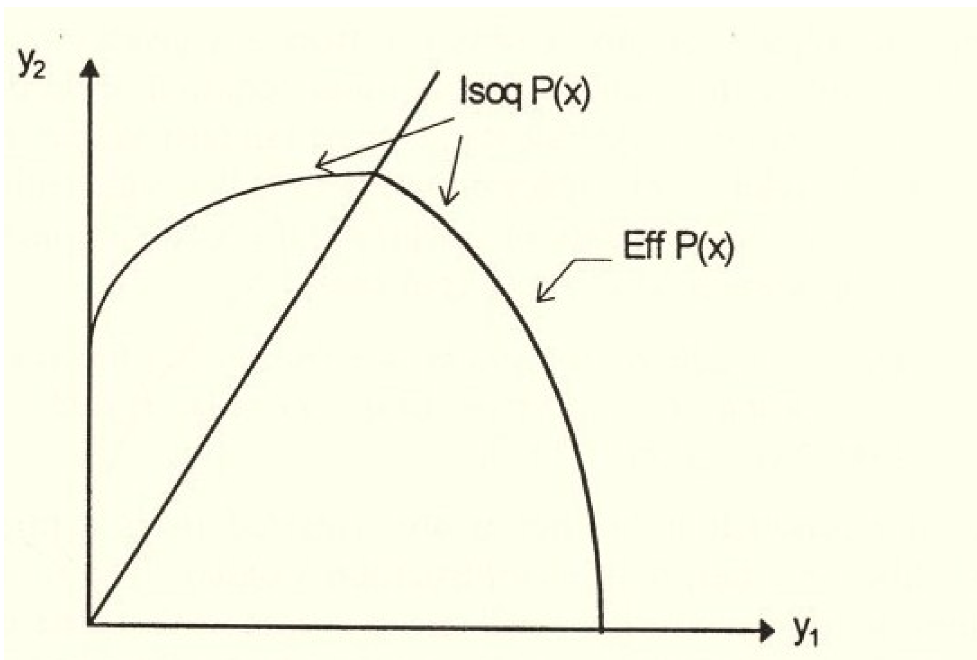

FIGURE 1.1.2. The output Isoquant and the Output Efficient Subset

In Figure 1.1.2, Eff $P(x)$ contains only the downward-sloping of Isoq $P(x)$. Eff $P(x)$ provides a more accurate way to measure the technical efficiency of output production since Eff $P(x) \subseteq$ Isoq $P(x)$. In the case that we have one output, Isoq $P(x)$ = Eff $P(x)$.

DEFINITION 1.1.5. *A production frontier is a function* $f(x) = max\,\{y : y \in P(x)\}$

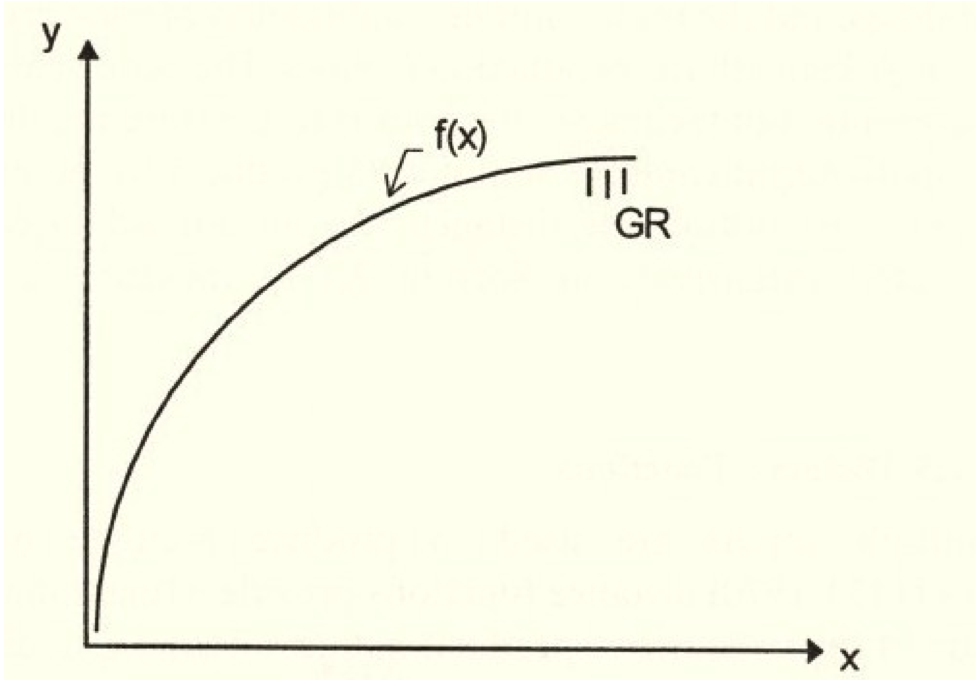

FIGURE 1.1.3. A Production Frontier

Figure 1.1.3 illustrates the production frontier $f(x)$ as the upper boundary of GR. The production frontier $f(x)$ describes the maximum output that can be produced with any given input.

The production frontier defines technical efficiency in terms of a maximum output produced by a given set of inputs. This approach involves selecting the mix of true inputs which produces a given quantity of output at a minimum cost, namely the production frontier. If what a producer actually produces is less than what it could feasibly produce, then it will lie below the frontier. The distance by which a firm lies below its production frontier is a measure of the firm's inefficiency (Bera and Sharma, 1996).

1.2. Technical Efficiency

A formal definition of technical efficiency is achieving "maximum output from a given input vector" (Koopmans, 1951).

Definition 1.2.1. *An output-input vector* $(y, x) \in GR$ *is technically efficient if, and only if,* $(y', x') \notin GR$ *for* $(y', -x') \geq (y, -x)$.

Definition 1.2.2. *An output vector* $y \in P(x)$ *is technically efficient if, and only if,* $y' \notin P(x)$ *for* $y' \geq y$ *or,* $y \in Eff\ P(x)$.

The following measure of technical efficiency was first proposed by Debreu (1951) and Farrell (1957), and they are often known as Debreu-Farrell measures of technical efficiency.

Definition 1.2.3. *An output-oriented measure of technical efficiency is a function* $TE(x, y) = [max\{\phi : \phi y \in P(x)\}]^{-1}$.

Kumbhakar and Lovell (2003) also gave the following result

Proposition 1.2.4. *The output-oriented measure of technical efficiency* $TE(x, y)$ *satisfies the properties:*

(i) $TE(x, y) \leq 1$

(ii) $TE(x, y) = 1 \Leftrightarrow y \in IsoqP(x)$.

TE is used to measure the technical efficiency. The first property is a normalization property, which states that TE is bounded above by one.

In Chapter 2, we present a method to estimate technical efficiency and explain how technical efficiency is related to stochastic frontier models. Then, three stochastic frontier model variations were introduced and the same six steps will have to be applied to analyse technical efficiency. In Chapter 3, we discuss and explain the distribution patterns of truncated skewed-Laplace. In Chapter 4, we propose the normal-truncated skewed-Laplace model in stochastic frontier analysis. In Chapter 5, we perform a simulation study to test normal-truncated skewed-Laplace and normal-exponential models. The purpose is to observe how close the true technical efficiency and the estimated technical efficiency are. In Chapter 6, we conclude that the normal-truncated skewed-Laplace model is more beneficial than normal-exponential model for stochastic frontier analysis particularly for the one data set used in this paper. Hence we suggest that further simulations should be conducted to test the effects of the normal-truncated skewed-Laplace model with more data sets.

Chapter 2

The Estimation of Technical Efficiency

We assume that cross-sectional data (observation on each producer at one point in time) on the quantities of N inputs used to produce a single output are available for each of I producers. A production frontier model without a random component can be written as

$$y_i = f(x_i, \beta) * TE_i, \tag{2.1}$$

where y_i is the observed scalar output of producer i (i=1,2,...I), x_i is a vector of N inputs used by producer i, $f(x_i;\beta)$ is the production frontier, and β is a vector of technology parameters to be estimated. TE_i denotes the technical efficiency defined as the ratio of observed output to maximum feasible output. TE_i can be expressed as

$$TE_i = \frac{y_i}{f(x_i, \beta)}. \tag{2.2}$$

Since TE_i is the output-oriented technical efficiency of producer i, the producer obtains the maximum feasible value of $y_i = f(x_i, \beta)$ if and only if $TE_i = 1$. Otherwise $TE_i < 1$ provides a measure of the shortfall of observed output from the maximum feasible output.

A stochastic component that describes random shocks affecting the production process is to be considered. We denote these effects with $exp\{v_i\}$. Each producer is facing a different shock, but we assume the shocks are random, and they are described by a common distribution. Thus, the stochastic production frontier for cross-sectional data becomes

$$y_i = f(x_i, \beta) \cdot exp\{v_i\} \cdot TE_i, \tag{2.3}$$

where $exp\{v_i\}$ represents the random shocks on each producer, and $f(x_i, \beta) \cdot exp\{v_i\}$ the stochastic production frontier. Then the output-oriented technical efficiency of producer i is

$$TE_i = \frac{y_i}{f(x_i, \beta) \cdot exp\{v_i\}}, \tag{2.4}$$

which defines technical efficiency as the ratio of observed output to the maximum feasible output, conditional on $exp\{v_i\}$. Producer i attains its maximum feasible output of $f(x_i, \beta) \cdot exp\{v_i\}$ if and only if $TE_i = 1$. Otherwise, $0 < TE_i < 1$ provides a measure of the shortfall of observed output from the maximum feasible output in an environment characterized by $exp\{v_i\}$ (see Kumbhakar and Lovell (2003), page 64-66).

The only difference between these models is that the deterministic models (2.1) and (2.2) ignore the random shocks during the production process, and model (2.3) and (2.4) include these effects. We are able to estimate the technical efficiency using stochastic production frontier model (2.3) and (2.4).

2.1. General Procedure to Estimate Technical Efficiency

Technical efficiency, TE_i, can be obtained as the exponential conditional expectation of u given the composed error term ϵ, which is given by

$$TE_i = e^{-E(u_i|\epsilon_i)}. \tag{2.5}$$

This estimation was proposed by Jondrow, Lovell, Materov, and Schmidt in 1982. The expectation of u_i given ϵ_i can be obtained from the joint density function of u_i and ϵ_i, and marginal density function of $f(\epsilon)$. The joint density function of u_i and ϵ_i is the product of individual density functions u and ϵ. The marginal density function $f(\epsilon)$ is obtained by integrating u out of $f(u, \epsilon)$.

Aigner and Chu (1968) were the first researchers to use the maximum likelihood estimate (MLE) method to estimate point estimators. The MLE method can be used to maximize the log likelihood function corresponding to the marginal density function, $f(\epsilon)$. Once the maximum point estimates of parameters are obtained, we substitute these values into (2.5) to estimate the technical efficiency of each producer.

2.2. Stochastic Production Frontier Models

Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977) simultaneously introduced stochastic production frontier models in the form of (2.3). They use $exp\{-u_i\}$ to present technical efficiency. Thus, the equation (2.3) becomes

$$y_i = f(x_i, \beta) \cdot exp\{v_i\} \cdot exp\{-u_i\}. \tag{2.6}$$

In order to estimate the stochastic frontier model, we assume that $f(x_i, \beta)$ take the log-linear Cobb-Douglas form; then equation (2.3) can be written as

$$\ln y_i = \beta_0 + \sum_n \beta_n \ln x_{ni} + v_i - u_i. \tag{2.7}$$

The error term in this equation, $\epsilon = v_i - u_i$, is composed of a two-sided "noise" component, v_i, and a nonnegative technical inefficiency component, u_i. The noise component v_i is assumed to be independent and identically distributed (iid) and symmetric, distributed independently of u_i. If $\epsilon_i > 0$, u_i is not large, which means that this producer is relatively efficient, while if $\epsilon_i < 0$, u_i is large, which means that this producer is relatively inefficient. The maximum output can be reached when $u_i = 0$.

2.2.1. The Normal Half-Normal Model. Kumbhaker and Knox Lovell (2003) were the first to comprehensively explain the normal half-normal model. In this model, they begin with three distributional assumptions:

(1) $v_i \sim iidN(0, \sigma_v^2)$;

(2) $u_i \sim iidN^{+}(0, \sigma_u^2)$, that is, as nonnegative half normal;

(3) u_i and v_i are distributed independently of each other and of the regressors. The density function of $u \geq 0$ is given in by

$$f(u) = \frac{2}{\sqrt{2\pi}\sigma_u} exp\left\{-\frac{u^2}{2\sigma_u^2}\right\}. \tag{2.8}$$

The density function of v is

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} exp\left\{-\frac{v^2}{2\sigma_v^2}\right\}, -\infty < v < \infty. \tag{2.9}$$

Assuming independence of the error terms v and u, the joint density function results as the product of individual density functions,

$$f(u, v) = \frac{2}{2\pi\sigma_u\sigma_v} exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right\}. \tag{2.10}$$

Since $\epsilon = v - u$, the joint density function for u and ϵ is

$$f(u, \epsilon) = \frac{2}{2\pi\sigma_u\sigma_v} exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{(\epsilon + u)^2}{2\sigma_v^2}\right\}. \tag{2.11}$$

The marginal density function of ϵ is obtained by integrating u out of $f(u, v)$, which gives

$$\begin{aligned}
f(\epsilon) &= \int_0^\infty f(u, \epsilon)\, du \\
&= \int_0^\infty \frac{2}{2\pi\sigma_u\sigma_v} exp\left\{-\frac{u^2}{2\sigma_u^2} - \frac{(\epsilon + u)^2}{2\sigma_v^2}\right\} du \\
&= \frac{2}{\sqrt{2\pi}\sigma} \cdot \left[1 - \Phi\left(\frac{\epsilon\lambda}{\sigma}\right)\right] exp\left\{-\frac{\epsilon^2}{2\sigma^2}\right\} \\
&= \frac{2}{\sigma} \cdot \phi\left(\frac{\epsilon}{\sigma}\right) \cdot \Phi\left(-\frac{\epsilon\lambda}{\sigma}\right),
\end{aligned} \tag{2.12}$$

where $\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2}$, $\lambda = \frac{\sigma_u}{\sigma_v}$, and $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal cumulative distribution and density functions, respectively. The marginal density function of ϵ is asymmetric and characterized by

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

The variance of ϵ is given by

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

Using (2.12), the log likelihood function for a sample of I producers is given by

$$\ln L(\epsilon|\lambda, \sigma^2) = n \ln\left(\frac{\sqrt{2}}{\sqrt{\pi}}\right) - n \ln \sigma + \sum_{i=1}^{n} \Phi\left(-\frac{\lambda \epsilon_i}{\sigma}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n} \epsilon_i^2. \tag{2.13}$$

The log likelihood function in (2.13) can be maximized with respect to the parameters to obtain maximum likelihood estimates of all parameters.

To obtain estimates of the technical efficiency of each producer, we have to estimate ϵ. This can be obtained from the conditional distribution of u_i given ϵ_i. Jondrow et al.(1982) show that if $u_i \sim N^+(0, \sigma_u^2)$, the conditional distribution of u given ϵ is

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

where $\mu_* = -\frac{\epsilon \sigma_u^2}{\sigma^2}$ and $\sigma_*^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma^2}$. We can see that the distribution of u conditional on ϵ is $N^+(\mu_*, \sigma_*^2)$, thus the expected value, $E(u|\epsilon)$, can be used as a point estimator for u. This is given by

$$E(u_i|\epsilon_i) = \mu_{*i} + \sigma_* \left[\frac{\phi(-\frac{\mu_{*i}}{\sigma_*})}{1 - \Phi(-\frac{\mu_{*i}}{\sigma_*})}\right]. \tag{2.14}$$

Once a point estimates of u_i are obtained, estimates of the technical efficiency of each producer can be obtained from (2.5).

2.2.2. The Normal Exponential Model. Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broeck (1977) were the first researchers to use the normal-exponential model. In this model, they assume three distributional assumptions:

(1) $v_i \sim$ iid $N(0, \sigma_v^2)$;

(2) $u_i \sim$ iid $Exp(\sigma_u)$;

(3) u_i and v_i are distributed independently of each other and of the regressors. To obtain the joint density function of u and v, we can multiply their individual density functions $f(v)$ form (2.9) and $f(u)$, given by

$$f(u) = \frac{1}{\sigma_u} \cdot exp\left\{-\frac{u}{\sigma_u}\right\}. \tag{2.15}$$

Thus, the joint density function is

$$f(u,v) = \frac{1}{\sqrt{2\pi}\sigma_u\sigma_v} \cdot exp\left\{-\frac{u}{\sigma_u} - \frac{v^2}{2\sigma_v^2}\right\}. \tag{2.16}$$

The joint density function of u and ϵ is given by

$$f(u,\epsilon) = \frac{1}{\sqrt{2\pi}\sigma_u\sigma_v} \cdot exp\left\{-\frac{u}{\sigma_u} - \frac{(u+\epsilon)^2}{2\sigma_v^2}\right\}. \tag{2.17}$$

Thus, the marginal density function of ϵ is

$$\begin{aligned} f(\epsilon) &= \int_0^\infty f(u,\epsilon)\,du \\ &= \frac{1}{\sigma_u} \cdot \Phi\left(-\frac{\epsilon}{\sigma_v} - \frac{\sigma_v}{\sigma_u}\right) \cdot exp\left\{\frac{\epsilon}{\sigma_u} + \frac{\sigma_v^2}{2\sigma_u^2}\right\}. \end{aligned} \tag{2.18}$$

The marginal density function of ϵ is asymmetric and characterized by

$$E(\epsilon) = -E(u) = -\sigma_u$$

and

$$V(\epsilon) = \sigma_u^2 + \sigma_v^2.$$

The log-likelihood function for a sample of I producers can be written as

$$\ln L(\epsilon|\sigma_u,\sigma_v) = n \ln\left(\frac{1}{\sqrt{2\pi}}\right) - n\ln\sigma_u + n\left(\frac{\sigma_v^2}{2\sigma_u^2}\right) + \sum_{i=1}^{n} \ln\Phi(-A) + \sum_{i=1}^{n}\frac{\epsilon_i}{\sigma_u}, \tag{2.19}$$

where $A = -\frac{\mu^+}{\sigma_v}$ and $\mu^+ = -\epsilon - \left(\frac{\sigma_v^2}{\sigma_u}\right)$. Equation (2.19) $\ln L$ can be maximized with respect to the parameters to obtain maximum likelihood estimates of all parameters.

The conditional distribution $f(u|\epsilon)$ is distributed as $N^+(\mu^+, \sigma_v^2)$, and the density function is given by

$$f(u|\epsilon) = \frac{f(u,\epsilon)}{f(\epsilon)} = \frac{1}{\sqrt{2\pi}\sigma_v\Phi(-\mu^+/\sigma_v)} \cdot exp\left\{-\frac{(u-\mu^+)^2}{2\sigma^2}\right\}. \tag{2.20}$$

The expected value of inefficiency term u given ϵ in the normal-exponential model is

$$E(u_i|\epsilon_i) = \mu_i^+ + \sigma_v\left[\frac{\phi(-\mu_i^+/\sigma_v)}{\Phi(\mu_i)^+/\sigma_v)}\right] = \sigma_v\left[\frac{\phi(A)}{\Phi(-A)} - A\right]. \tag{2.21}$$

2.2.3. The Normal-Truncated Normal Model. The normal-truncated normal model was introduced by Stevenson (1980). In this model, three distributional assumptions are as follow:

(1) $v_i \sim iidN(0, \sigma_v^2)$;

(2) $u_i \sim iidN^+(\mu, \sigma_u^2)$;

(3) u_i and v_i are distributed independently of each other and of the regressors. The density function $f(v)$ is given in equation (2.9) and truncated normal density function for $u \geq 0$ is given by

$$f(u) = \frac{1}{\sqrt{2\pi}\sigma_u\Phi(\mu/\sigma_u)} \cdot exp\left\{-\frac{(u-\mu)^2}{2\sigma_u^2}\right\}, \tag{2.22}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, and μ is normally distributed, which is truncated below at zero.

The joint density function of u and v is

$$f(u,v) = \frac{1}{2\pi\sigma_u\sigma_v\Phi(\mu/\sigma_u)} \cdot exp\left\{-\frac{(u-\mu)^2}{2\sigma_u^2} - \frac{v^2}{2\sigma_v^2}\right\}. \tag{2.23}$$

The joint density of u and ϵ is

$$f(u,\epsilon) = \frac{1}{2\pi\sigma_u\sigma_v\Phi(\mu/\sigma_u)} \cdot exp\left\{-\frac{(u-\mu)^2}{2\sigma_u^2} - \frac{(\epsilon+u)^2}{2\sigma_v^2}\right\}. \tag{2.24}$$

The marginal density of ϵ is

$$
\begin{aligned}
f(\epsilon) &= \int_0^\infty f(u,\epsilon)\,du \\
&= \frac{1}{\sqrt{2\pi}\sigma\Phi(\mu/\sigma_u)} \cdot \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\epsilon\lambda}{\sigma}\right) \cdot exp\left\{-\frac{(\epsilon+\mu)^2}{2\sigma^2}\right\} \\
&= \frac{1}{\sigma} \cdot \phi\left(\frac{\epsilon+\mu}{\sigma}\right) \cdot \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\epsilon\lambda}{\sigma}\right) \cdot \left[\Phi\left(\frac{\mu}{\sigma_u}\right)\right]^{-1},
\end{aligned}
$$

where $\sigma = (\sigma_u^2 + \sigma_v^2)^{1/2}$ and $\lambda = \frac{\sigma_u}{\sigma_v}$, and $\phi(\cdot)$ is the standard normal density function. The marginal density function of ϵ is asymmetric and characterized by

$$
E(\epsilon) = -E(\epsilon) = -\frac{\mu a}{2} - \frac{\sigma_u a}{\sqrt{2\pi}} \cdot exp\left\{-\frac{1}{2}\left(\frac{\mu}{\sigma_u}\right)^2\right\}
$$

and the variance is

$$
V(\epsilon) = \mu^2 \frac{a}{2}\left(1 - \frac{a}{2}\right) + \frac{a}{2}\left(\frac{\pi - a}{\pi}\right)\sigma_u^2 + \sigma_v^2.
$$

The log likelihood function is

$$
\ln L = n \ln \frac{1}{\sqrt{2\pi}} - n \ln \sigma - nln\Phi\left(\frac{\mu}{\sigma_u}\right) + \sum_{i=1}^{n} \ln \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\epsilon_i\lambda}{\sigma}\right) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\epsilon_i + \mu}{\sigma}\right)^2, \tag{2.25}
$$

where $\sigma_u = \lambda\sigma/\sqrt{1+\lambda^2}$.

The conditional distribution $f(u|\epsilon)$ is distributed as $N^+(\mu_i^+, \sigma_*^2)$, given by

$$
f(u|\epsilon) = \frac{f(u,\epsilon)}{f(\epsilon)} = \frac{1}{\sqrt{2\pi}\sigma_*\left[1 - \Phi(-\mu^+/\sigma_*)\right]} \cdot exp\left\{-\frac{(u-\mu^+)^2}{2\sigma_*^2}\right\}, \tag{2.26}
$$

where $\mu^+ = (-\sigma_u^2\epsilon_i + \mu\sigma_v^2)/\sigma^2$ and $\sigma_*^2 = \sigma_u^2\sigma_v^2/\sigma^2$. The expected value of inefficiency u, given estimated ϵ, in the normal-exponential model is

$$
E(u_i|\epsilon_i) = \sigma_*\left[\frac{\mu_i^+}{\sigma_*} + \frac{\phi(\mu_i^+/\sigma_*)}{1 - \Phi(-\mu_i^+/\sigma_*)}\right]. \tag{2.27}
$$

In this paper, we will consider a normal-truncated skewed-Laplace model which combines the normal distribution and the truncated skewed-Laplace distribution. In the next chapter, we will introduce the truncated skewed-Laplace distribution.

Chapter 3

Truncated Skewed-Laplace Distribution

3.1. Truncated Skewed-Laplace Distribution

The Truncated skewed-Laplace distribution (TSL) is a generalized model of exponential distribution. The TSL distribution provides more flexible representation of the pattern of efficiency in the data. This model was first introduced by Aryal and Rao in 2005. The standard Laplace distribution has probability density function (pdf) given by G(X) given by

$$g(x) = \frac{1}{2\phi} exp\left(-\frac{|x|}{\phi}\right) \tag{3.1}$$

and cumulative distribution function (cdf)

$$G(x) = \begin{cases} \frac{1}{2} exp\left(\frac{x}{\phi}\right) & \text{if } x \leq 0 \\ 1 - \frac{1}{2} exp\left(-\frac{x}{\phi}\right) & \text{if } x \geq 0, \end{cases} \tag{3.2}$$

where $-\infty < x < \infty$ and $\phi > 0$. A random variable X has the skew-Laplace distribution if its pdf is

$$f(x) = 2g(x)G(\lambda x), \tag{3.3}$$

where $x \in \mathbb{R}$ and $\lambda \in \mathbb{R}$. According to (3.3), the pdf and the cdf of X are

$$f(x) = \begin{cases} \frac{1}{2\phi} exp\left\{-\frac{(1+|\lambda|)|x|}{\phi}\right\} & \text{if } \lambda x \leq 0 \\ \frac{1}{\phi} exp\left(-\frac{|x|}{\phi}\right)\left\{1 - \frac{1}{2} exp\left(-\frac{\lambda x}{\phi}\right)\right\} & \text{if } \lambda x > 0 \end{cases} \tag{3.4}$$

and

$$F(x) = \begin{cases} \frac{1}{2} + \frac{sign(\lambda)}{2}\left[\frac{1}{1+|\lambda|} exp\left\{-\frac{(1+|\lambda|)|x|}{\phi}\right\} - 1\right] & \text{if } \lambda x \leq 0 \\ \frac{1}{2} + sign(\lambda)\left[\frac{1}{2} - exp\left(-\frac{|x|}{\phi}\right) \times \left\{1 - \frac{1}{2(1+|\lambda|)} exp\left(-\frac{\lambda x}{\phi}\right)\right\}\right] & \text{if } \lambda x > 0, \end{cases} \tag{3.5}$$

respectively, where λ is a skewness parameter.

In this study, we consider the case when the skew-Laplace distribution is truncated on the left at zero, and we assume that $\lambda > 0$. The cdf of the truncated skewed Laplace

(TSL) distribution is given by

$$F^*(x) = 1 + \frac{exp\left(-\frac{(1+\lambda)x}{\phi}\right) - 2(1+\lambda)exp(-\frac{x}{\phi})}{(2\lambda+1)}, \tag{3.6}$$

and the corresponding pdf is

$$f^*(x) = \frac{(1+\lambda)}{\phi(2\lambda+1)}\left\{2exp\left(-\frac{x}{\phi}\right) - exp\left(-\frac{(1+\lambda)x}{\phi}\right)\right\}, x > 0. \tag{3.7}$$

When $\lambda = 0$, the density of TSL becomes the density of the exponential distribution.

3.2. Moments

The kth moment of X, based on the definition of the gamma function, is given by

$$E(X^k) = \frac{\phi^k(1+\lambda)\Gamma(k+1)}{(2\lambda+1)}\left\{2 - \frac{1}{(1+\lambda)^{k+1}}\right\}, \tag{3.8}$$

The proof of this is seen by the following equations.

Proof.

$$\begin{aligned}
E(X^k) &= \int_0^\infty x^k \cdot f^*(x)\,dx \\
&= \frac{(1+\lambda)}{\phi(2\lambda+1)}\left[\int_0^\infty 2x^k exp\left(-\frac{x}{\phi}\right)dx - \int_0^\infty x^k exp\left(-\frac{(1+\lambda)x}{\phi}\right)dx\right] \\
&= \frac{(1+\lambda)}{\phi(2\lambda+1)}\left[2\phi^{k+1}\Gamma(k+1) - \frac{\phi^{k+1}}{(1+\lambda)^{k+1}}\Gamma(k+1)\right] \\
&= \frac{\phi^k(1+\lambda)\Gamma(k+1)}{(2\lambda+1)}\left[2 - \frac{1}{(1+\lambda)^{k+1}}\right].
\end{aligned}$$

□

Using the binomial expansion and (3.8), the kth central moment of X can be calculated as

$$E\left[(x-\mu)^k\right] = \begin{cases} = \mu^k + \sum_{j=1}^{k/2} \begin{pmatrix} k \\ 2j \end{pmatrix} \mu^{k-2j} \phi^{2j} \frac{(1+\lambda)\Gamma(2j+1)}{(2\lambda+1)} \\ \times \left\{2 - \frac{1}{(1+\lambda)^{2j+1}}\right\} - \sum_{j=1}^{k/2} \begin{pmatrix} k \\ 2j-1 \end{pmatrix} \mu^{k-2j+1} \phi^{2j-1} \\ \times \frac{(1+\lambda)\Gamma(2j)}{(2\lambda+1)} \left\{2 - \frac{1}{(1+\lambda)^{2j}}\right\}, \; if \; k \; is \; even \\ = -\mu^k - \sum_{j=1}^{(k-1)/2} \begin{pmatrix} k \\ 2j \end{pmatrix} \mu^{k-2j} \phi^{2j} \frac{(1+\lambda)\Gamma(2j+1)}{(2\lambda+1)} \\ \times \left\{2 - \frac{1}{(1+\lambda)^{2j+1}}\right\} + \sum_{j=1}^{(k+1)/2} \begin{pmatrix} k \\ 2j-1 \end{pmatrix} \mu^{k-2j+1} \phi^{2j-1} \\ \times \frac{(1+\lambda)\Gamma(2j)}{(2\lambda+1)} \left\{2 - \frac{1}{(1+\lambda)^{2j}}\right\}, \; if \; k \; is \; odd \end{cases}$$

where $\mu = E(X)$ is the expectation of X. It follows from the above equations that the expectation and variance of X are

$$E(X) = \phi \frac{(1+4\lambda+2\lambda^2)}{(1+\lambda)(1+2\lambda)} \tag{3.9}$$

and

$$Var(X) = \phi^2 \frac{(1+8\lambda+16\lambda^2+12\lambda^3+4\lambda^4)}{(1+\lambda)^2(1+2\lambda)^2} \tag{3.10}$$

If $\lambda = 0$, the mean and variance of X are the mean and variance of the $Exp(\phi)$ distribution.

3.3. Skewness

The skewness coefficient is a measure of the asymmetry of the probability distribution of a real-valued random variable. The skewness coefficient can be positive or negative, or even undefined. Negative values for the skewness coefficient indicate data that are left skewed, and positive values for the skewness coefficient indicate data that are right skewed. A left skewed distribution has a longer left tail than right tail, and the mass of the distribution is concentrated on the right of the distribution. Similarly, a right skewed distribution has a longer right tail than left tail, and the mass of the distribution is concentrated on the left of the distribution.

The skewness of a random variable X is the third standardized moment, and is denoted by γ_1. Skewness is defined as

$$\begin{aligned}\gamma_1 &= E\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] \\ &= \frac{E\left[X^3\right]-3\mu E\left[X^2\right]+3\mu^2 E\left[X\right]-\mu^3}{\sigma^3} \\ &= \frac{E\left[X^3\right]-3\mu\sigma^2-\mu^3}{\sigma^3}.\end{aligned}$$

For TSL,

$$\begin{aligned}\gamma_1 = & \frac{1}{(1+8\lambda+16\lambda^2+12\lambda^3+4\lambda^4)^{3/2}} \times \big[12(1+\lambda)^4(1+2\lambda)^2-6(1+2\lambda)^2 \\ & -3(1+4\lambda+2\lambda^2)(1+8\lambda+16\lambda^2+12\lambda^3+4\lambda^4)-(1+4\lambda+2\lambda^2)^3\big]\end{aligned}$$

since $E\left[X^2\right] = \frac{4\phi^2(1+\lambda)^3-2\phi^2}{(2\lambda+1)(1+\lambda)^2}$ and $E\left[X^3\right] = \frac{12\phi^3(1+\lambda)^4-6\phi^3}{(2\lambda+1)(1+\lambda)^3}$. When $\lambda = 0$, γ_1 is the exponential distribution skewness, which is 2.

Figure 3.3.1 shows that TSL skewness is unbounded below at the left side of the origin point, but as λ approaches to positive infinity the skewness coefficient is approaching 2.

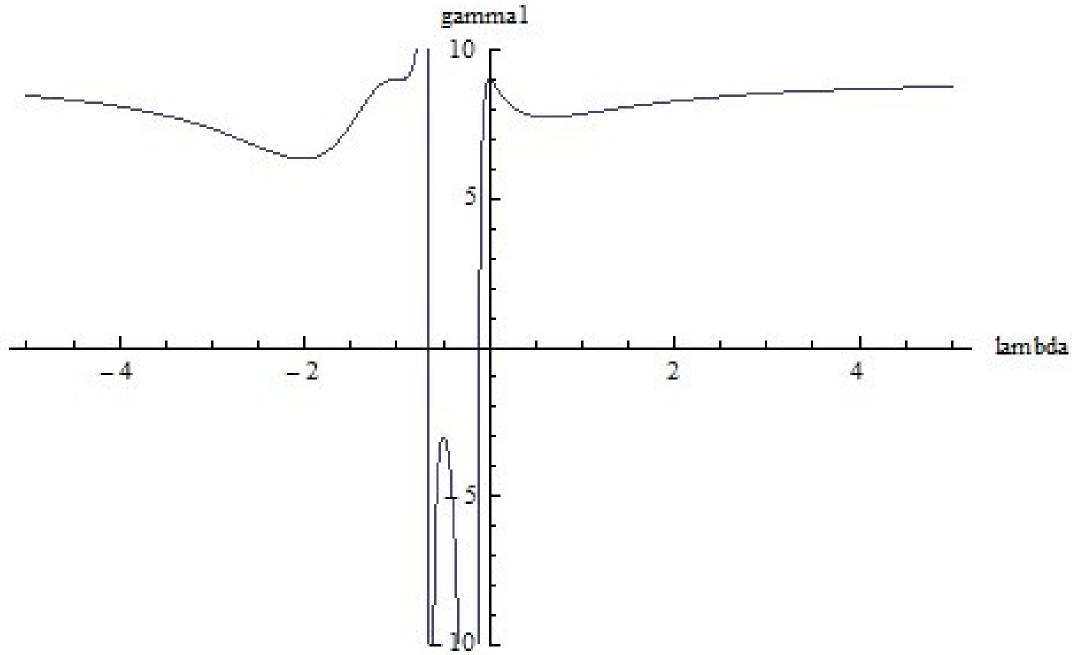

FIGURE 3.3.1. The Skewness of TSL

3.4. Kurtosis

Kurtosis is a measure of whether the data is flat or peaked. The kurtosis of a random variable X is the fourth standardized moment, denoted by γ_2 and defined as

$$\begin{aligned}
\gamma_2 &= E\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] \\
&= \frac{E\left(X^4\right)-4E\left(X^3\right)\mu+6E\left(X^2\right)\mu^2-4E\left(X\right)\mu^3+\mu^4}{\sigma^4} \\
&= \frac{E\left(X^4\right)-4E\left(X^3\right)+6E\left(X^2\right)\mu^2-3\mu^4}{\sigma^4}.
\end{aligned}$$

For TSL,

$$\gamma_2 = \frac{1}{\left(1+8\lambda+16\lambda^2+12\lambda^3+4\lambda^4\right)^2} \times \Big[48(1+\lambda)^5(1+2\lambda)^3 - 24(1+2\lambda)^3$$
$$-48(1+\lambda)^4(1+2\lambda)^2\left(1+4\lambda+2\lambda^2\right) + 24(1+2\lambda)^2\left(1+4\lambda+2\lambda^2\right)$$
$$+24(1+\lambda)^3(1+2\lambda)\left(1+2\lambda+2\lambda^2\right)^2 - 12(1+2\lambda)\left(1+4\lambda+2\lambda^2\right)^2$$
$$-3\left(1+2\lambda+2\lambda^2\right)^4\Big],$$

since $E\left(X^4\right) = \frac{48\phi^4(1+\lambda)^5 - 24\phi^4}{(1+\lambda)^4(1+2\lambda)}$. When $\lambda = 0$, γ_2 is the exponential distribution kurtosis, which is 9.

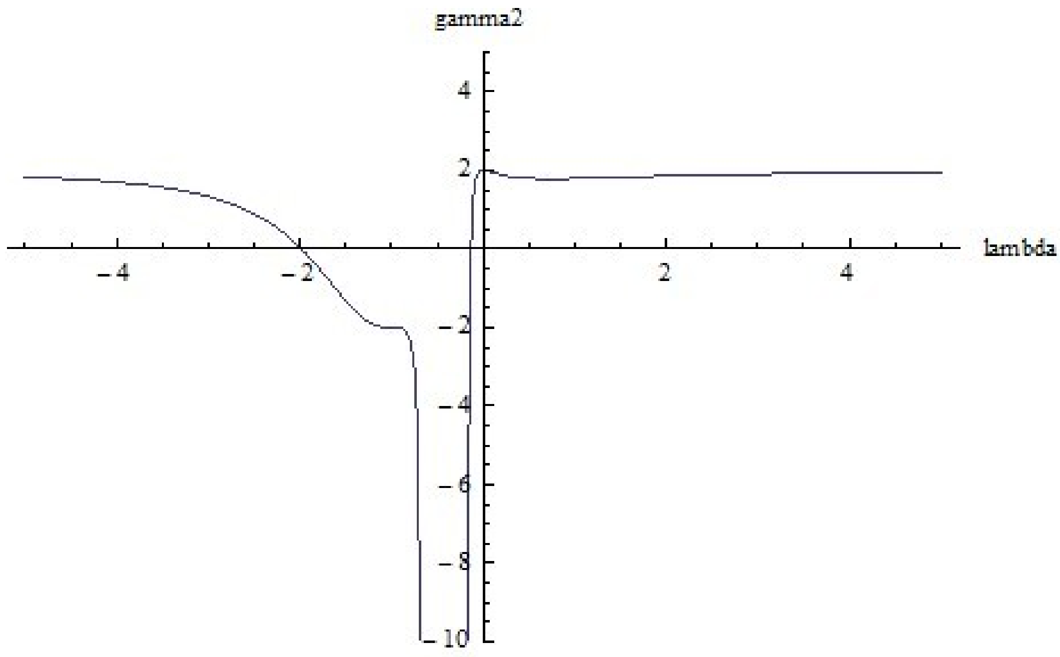

Figure 3.4.1. The Kurtosis of TSL

Figure 3.4.1 shows that the kurtosis of TSL is unbounded below at the left side of the origin point. The coefficient of kurtosis is bounded by 9 as λ approaches positive infinity from the origin point.

Chapter 4

The Normal-Truncated Skewed-Laplace in Stochastic Frontier Analysis

In this chapter, we propose a new model in stochastic frontier analysis, the normal-truncated skewed-Laplace. The normal-truncated skewed-Laplace (NTSL) model is a generalized model of the normal-exponential model when $\lambda = 0$. Compared with the normal-exponential model, NTSL is more flexible because it has a skewness parameter λ. In this model, we assume the following:

(1) $v_i \sim iidN(0, \sigma_v^2)$;

(2) $u_i \sim$ iid truncated skewed-Laplace;

(3) u_i and v_i are distributed independently of each other and of the regressors. The density of v is given by

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{v^2}{2\sigma^2}\right\}, v \in \mathbb{R}, \sigma > 0 \tag{4.1}$$

The density of u is

$$f(u) = \frac{(1+\lambda)}{\phi(2\lambda+1)} \left\{2exp\left(-\frac{u}{\phi}\right) - exp\left(-\frac{(1+\lambda)u}{\phi}\right)\right\}, u > 0, \lambda > 0, \phi > 0 \tag{4.2}$$

Given the independence assumption, the joint density function of u and v is the product of their individual density functions. Thus, the joint density function of u and v is

$$f(u,v) = \frac{(1+\lambda)}{\phi(2\lambda+1)} \cdot \left\{2exp\left(-\frac{v^2}{2\sigma^2} - \frac{u}{\phi}\right) - exp\left(-\frac{v^2}{2\sigma^2} - \frac{(1+\lambda)u}{\phi}\right)\right\} \tag{4.3}$$

If we substitute v=ϵ+u into $f(u,v)$, then the joint density function for u and ϵ becomes

$$f(u,\epsilon) = \frac{(1+\lambda)}{\sqrt{2\pi}\sigma\phi(2\lambda+1)} \cdot \left\{2exp\left(-\frac{(\epsilon+u)^2}{2\sigma^2} - \frac{u}{\phi}\right) - exp\left(-\frac{(\epsilon+u)^2}{2\sigma^2} - \frac{(1+\lambda)u}{\phi}\right)\right\}. \tag{4.4}$$

Thus, we can get the marginal density function of $f(\epsilon)$,

$$\begin{aligned}
f(\epsilon) &= \int_0^\infty f(u,\epsilon)\,du \\
&= \frac{2(1+\lambda)}{\sqrt{2\pi}\sigma\phi(2\lambda+1)} \cdot exp\left(\frac{\epsilon}{\phi}+\frac{\sigma^2}{2\phi^2}\right)\int_0^\infty exp\left[-\frac{1}{2}\left(\frac{u}{\sigma}+\frac{\epsilon}{\sigma}+\frac{\sigma}{\phi}\right)^2\right]du \\
&\quad -\frac{(1+\lambda)}{\sqrt{2\pi}\sigma\phi(2\lambda+1)} \cdot exp\left(\frac{\epsilon(1+\lambda)}{\phi}+\frac{(1+\lambda)^2\sigma^2}{2\phi^2}\right) \\
&\quad \cdot \int_0^\infty exp\left[-\frac{1}{2}\left(\frac{u}{\sigma}+\frac{\epsilon}{\sigma}+\frac{(1+\lambda)}{\phi}\right)^2\right]du \\
&= \frac{2(1+\lambda)}{\sqrt{2\pi}\sigma\phi(2\lambda+1)} \cdot exp\left(\frac{\epsilon}{\phi}+\frac{\sigma^2}{2\phi^2}\right)\int_{\frac{\epsilon}{\sigma}+\frac{\sigma}{\phi}}^\infty exp\left(\frac{z^2}{2}\right)\sigma\,dz \\
&\quad -\frac{(1+\lambda)}{\sqrt{2\pi}\sigma\phi(2\lambda+1)} \cdot exp\left(\frac{\epsilon(1+\lambda)}{\phi}+\frac{(1+\lambda)^2\sigma^2}{2\phi^2}\right)\int_{\frac{\epsilon}{\sigma}+\frac{(1+\lambda)\sigma}{\phi}}^\infty exp\left(\frac{z^2}{2}\right)\sigma\,dz \\
&= \frac{2(1+\lambda)}{\phi(2\lambda+1)}exp\left(\frac{\epsilon}{\phi}+\frac{\sigma^2}{2\phi^2}\right)\Phi\left(-\frac{\epsilon}{\sigma}-\frac{\sigma}{\phi}\right) \\
&\quad -\frac{(1+\lambda)}{\phi(2\lambda+1)}exp\left(\frac{\epsilon(1+\lambda)}{\phi}+\frac{(1+\lambda)^2\sigma^2}{2\phi^2}\right)\Phi\left(-\frac{\epsilon}{\sigma}-\frac{(1+\lambda)\sigma}{\phi}\right).
\end{aligned} \tag{4.5}$$

When λ=0, then the above equation becomes the marginal density of $f(\epsilon)$ in the normal-exponential model (2.18),

$$f(\epsilon) = \frac{1}{\phi}exp\left(\frac{\epsilon}{\phi}+\frac{\sigma^2}{2\phi^2}\right)\Phi\left(\frac{\epsilon}{\sigma}-\frac{\sigma}{\phi}\right).$$

The log likelihood function can be written as

$$\begin{aligned}
\ln L(\epsilon) &= \sum_{i=1}^N \ln\left[\frac{2\,(1+\lambda)}{\phi\,(1+2\lambda)}exp\left(\frac{\epsilon}{\phi}+\frac{\sigma^2}{2\phi^2}\right)\Phi\left(-\frac{\epsilon}{\sigma}-\frac{\sigma}{\phi}\right)\right. \\
&\quad \left.-\frac{(1+\lambda)}{\phi\,(1+2\lambda)}exp\left(\frac{\epsilon\,(1+\lambda)}{\phi}+\frac{\sigma^2\,(1+\lambda)^2}{2\phi^2}\right)\Phi\left(-\frac{\epsilon}{\sigma}-\frac{\sigma\,(1+\lambda)}{\phi}\right)\right].
\end{aligned} \tag{4.6}$$

The expected value of inefficiency u given ϵ in the normal-TSL model is

$$
\begin{aligned}
E(u|\epsilon) &= \frac{\int_0^\infty uf(u,\epsilon)\,du}{f(\epsilon)} \\
&= \frac{1}{2\sqrt{2\pi}exp\left(\frac{\epsilon}{\phi}+\frac{\sigma^2}{2\phi^2}\right)\Phi\left(-\frac{\epsilon}{\sigma}-\frac{\sigma}{\phi}\right)-\sqrt{2\pi}exp\left(\frac{\epsilon(1+\lambda)}{\phi}+\frac{(1+\lambda)^2}{2\phi^2}\right)\Phi\left(-\frac{\epsilon}{\sigma}-\frac{\sigma}{\phi}\right)} \\
&\quad \times\left[\sigma exp\left(-\frac{\epsilon}{2\sigma^2}\right)-2\sqrt{2\pi}\left(\frac{\sigma^2}{\phi}+\epsilon\right)exp\left(\frac{\epsilon}{\phi}+\frac{\sigma^2}{2\phi^2}\right)\Phi\left(-\frac{\epsilon}{\sigma}-\frac{\sigma}{\phi}\right)\right. \\
&\quad +\sqrt{2\pi}\left(\frac{(1+\lambda)\sigma^2}{\phi}+\epsilon\right)exp\left(\frac{(1+\lambda)\epsilon}{\phi}+\frac{(1+\lambda)^2\sigma^2}{2\phi^2}\right) \\
&\quad \left.\times\,\Phi\left(\frac{\epsilon}{\sigma}-\frac{(1+\lambda)\sigma}{\phi}\right)\right]. \qquad (4.7)
\end{aligned}
$$

Chapter 5

Simulation

In this chapter, we apply the normal-truncated skewed Laplace model and normal-exponential model to estimate technical efficiency. To assess the performance of the efficiency score estimation, we calculate the mean square error

$$MSE = \frac{1}{I}\sum_{i=1}^{I}\left(\widehat{TE_i} - TE_i\right)^2$$

between true and estimated efficiency scores for I producers. Additionally, we want to compare the normal-truncated skewed-Laplace and the normal-exponential using the MSE.

In Chapter 2, (2.4) is used to estimate the true technical efficiency, this can be simplified to $TE_i = exp(-u_i)$. We apply $\widehat{TE_i} = exp\left[-E(u_i|\epsilon_i)\right]$ to obtain estimates of individual efficiencies.

5.1. Normal-Truncated Skewed Laplace Simulation

In this section, we apply the normal-truncated skewed-Laplace model to estimate individual technical efficiency scores for simulated data based on a sample size of 100. We set the starting values of each point estimators at $\sigma = 0.1$, $\lambda = 0.03$, $\phi = 0.05$, $b_0 = 3$, $b_1 = 0.5$, $b_2 = 0.25$. The input vectors x_1 and x_2 are drawn independently from $N(10, 2)$ and $N(2, 0.2)$ distributions, respectively. The error term v_i follows the $N(0, \sigma)$ distribution. The data sets are included in Appendix 1.

There is no such command in R program that can simulate the parameters in truncated-skew Laplace distribution, thus a particular method needs to be used to simulate the inefficiency term u_i which is inverse transformation method (Luc Devroye, (1986)). The inverse transformation method states that if X is a random variable with cumulative distribution function F_X, then the random variable $Y = F_X(X)$ has a uniform$(0,1)$ distribution. According to this method, we simulate 100 random variables from uniform distribution over $(0,1)$. Using R program, we have $t = runif(100,0,1)$. The truncated skewed-Laplace cumulative distribution function can be computed using (3.6). Thus, we have

$$F^*(x) = 1 + \frac{exp\left(-\frac{(1+\lambda)x}{\phi}\right) - 2(1+\lambda)exp(-\frac{x}{\phi})}{(2\lambda+1)} = t.$$

In order to obtain u_i, we need to rewrite the function as

$$Y = 1 + \frac{exp\left(-\frac{(1+\lambda)u}{\phi}\right) - 2(1+\lambda)exp(-\frac{u}{\phi})}{(2\lambda+1)} - t = 0.$$

We use the *uniroot* command to numerically solve the above equation to get values of u_i

$$u_i = uniroot(Y, c(0,200), lambda = lambda, phi = phi, t = t[i])\$root,$$

where $c(0,200)$ is a vector containing the end-points of the interval to be searched for the root (see appendix 3). Once the values of u_i are obtained, we substitute these values into equation $TE_i = exp(-u_i)$ to get the true technical efficiency for each producer. The true values of technical efficiency are included in Appendix 1. For estimated technical efficiency, we need to obtain the estimated values of each parameter, and then substitute these values into $\widehat{TE_i} = exp\left[-E(u_i|\epsilon_i)\right]$ to get $\widehat{TE_i}$.

Optimization using the “optim” function in R is the first method we considered for calculating the estimated technical efficiency corresponding to the log likelihood function of $f(\epsilon)$. However, the optim method does not give us a desirable result because it is stuck at a local maximum, which is excessively small. An alternative method is used to calculate technical efficiency. We search for the maximum value of the log likelihood function manually using (4.6). We expect the estimated parameters to be close to the true values on either side, hence we determine the searching interval for each parameter. Next, we create six for-loops to locate the values of parameters in their intervals with small increments which result in the maximum value of the log likelihood function (see Appendices 4, 5, and 6). The estimated values of each parameter are $\widehat{\sigma} = 0.09$, $\widehat{\lambda} = 0.06$, $\widehat{\phi} = 0.065$, $\widehat{b_0} = 2.91$, $\widehat{b_1} = 0.52$, and $\widehat{b_2} = 0.27$. We substitute these values into the equation . The maximum likelihood value of the likelihood function is 77.96915. The MSE between true TE and estimated $\widehat{TE}$ is 0.046505867.

5.2. Normal-Exponential Model Simulation

We use the same approach to estimate technical efficiency using the normal-exponential model (see Appendices 7, 8, and 9). The estimated values of the parameters are $\widehat{\sigma_u} = 0.0675$, $\widehat{\sigma_v} = 0.09$, $\widehat{b_0} = 2.91$, $\widehat{b_1} = 0.52$, and $\widehat{b_2} = 0.27$ and the maximum likelihood value is 77.95368. The MSE between true TE and estimated $\widehat{TE}$ is 0.12798839.

In summary, the MSE of normal-truncated skewed Laplace model is less than the MSE of normal-exponential model, which is expected because we simulate the data according to normal-TSL model. In this paper, normal-truncated skewed-Laplace distribution is better than normal-exponential distribution. This simulation study helps us to test the application of both models with one data set.

Chapter 6

Conclusion and Future Research

Econometricians have been developing stochastic frontier models since 1977. As a result, the models have taken on a rich variety of forms. They have applied the models to a number of areas, such as banks, hospitals, schools, industries, and so on. In this thesis, we create a new model which combines the normal distribution and truncated skewed-Laplace distribution, namely, the normal-truncated skewed-Laplace model. We use R commands to find the point estimates of parameters via the maximum likelihood estimate method. Once point estimates are obtained, the estimated technical efficiency is attained. The true technical efficiency is based on the assumed starting values. We use MSE values to observe how close the true technical efficiency and estimated technical efficiency are. The smaller the MSE value, the better the model is. In this thesis, the MSE value of normal-truncated skewed-Laplace model is less than the MSE value of normal-exponential model. This shows that the normal-truncated skewed-Laplace model is better than the normal-exponential model for this data. In practice, normal-truncated skewed-Laplace distribution provides more flexibilities in modeling data, and hence it provides an alternative model for economics in estimating technical efficiency.

An interesting direction for future research is to investigate technical efficiencies upon different starting values of each parameter. For example, we may simulate 200 data sets to obtain an estimated technical efficiency and true technical efficiency. Then we compare the technical efficiencies of the normal-exponential model and that of the normal-truncated skewed-Laplace model. We may apply the models to real data sets to get the estimated technical efficiencies for producers. According to the value of the log likelihood function of each model, we can conclude which model is better. Moreover, other distributions could be considered, such as the normal-truncated gamma distribution, and so on.

APPENDIX

1. Tables

Table 1: True technical efficiency scores under both models

Observed output y	Input x_1	Input x_2	v	Random noise u	True TE
10.33872	8.303286	2.255659	0.078107	0.10251	0.902569
10.02234	10.09319	1.7513	-0.08616	0.003656	0.99635
9.404747	10.44442	1.919913	-0.1856	0.007901	0.99213
11.8205	10.43823	2.35184	-0.01193	0.003381	0.996625
11.55075	12.38001	1.976618	-0.01049	0.069757	0.93262
13.47615	13.45489	2.18816	0.100819	0.093947	0.910331
9.464449	10.85978	1.942944	-0.145	0.064656	0.937389
10.92637	8.906281	2.116586	0.024017	0.01228	0.987795
8.56031	8.173097	2.060451	-0.17283	0.009804	0.990244
10.85895	9.164571	1.991674	0.019512	0.013051	0.987034
17.06569	13.40372	2.016829	0.274179	0.008869	0.99117
8.505262	8.887892	2.394707	-0.26814	0.000447	0.999553
10.00758	9.090588	2.147516	-0.08951	0.000455	0.999545
10.34649	12.222	2.046547	-0.10425	0.088377	0.915416
11.98801	10.88915	2.098497	0.023587	0.017481	0.982671
13.7348	11.83895	2.236793	0.100969	0.016607	0.983531
11.30287	8.90564	1.843758	0.138993	0.058843	0.942855
10.26602	9.365772	1.919019	-0.02854	0.022717	0.977539
12.26879	9.727708	1.737798	0.175923	0.04312	0.957796
12.08117	10.92701	2.257344	0.012038	0.018169	0.981995
10.49746	8.170947	1.9489	0.055384	0.019972	0.980227
9.020327	8.337306	1.863757	-0.07395	0.041203	0.959634
11.86206	9.992102	2.22803	0.047697	0.024141	0.976148
12.25016	11.57331	1.887822	0.026633	0.002913	0.997091
8.278605	6.225581	1.704758	0.007714	0.040341	0.960462
11.85534	11.66207	1.775228	0.018674	0.016161	0.983969
12.96142	13.85057	1.660889	0.024123	0.001759	0.998243
8.879772	10.62551	1.936188	-0.12998	0.131661	0.876638
11.06605	12.03857	2.175552	-0.02249	0.11062	0.895279
11.32447	11.56077	2.224126	-0.00045	0.094847	0.909512
10.77449	11.39501	1.927789	-0.05577	0.046338	0.954719
10.03509	8.834633	2.231916	-0.06163	0.020946	0.979272
13.83064	12.22871	1.983103	0.108961	0.003746	0.996261
12.16369	12.94586	1.898207	-0.03674	0.004033	0.995975
10.47051	12.6565	1.781566	-0.15454	0.008967	0.991073
11.49886	8.318081	1.988612	0.127121	0.014561	0.985545
8.746989	6.956387	2.124934	0.013835	0.102003	0.903027
11.84318	10.85399	2.243619	0.111033	0.132182	0.876181
11.23536	11.89586	1.715104	-0.02423	0.028277	0.97212

Continued on next page

Table 1 – continued from previous page

Observed output y	Input x_1	Input x_2	v	Random noise u	True TE
11.15159	10.2413	1.934418	0.006748	0.021944	0.978295
13.01126	10.8067	2.141828	0.101498	0.014793	0.985316
12.77187	12.98239	2.022266	0.01565	0.024869	0.975438
9.058113	9.41839	1.966126	-0.18296	0.002339	0.997664
9.922498	6.963869	1.939268	0.149133	0.088886	0.91495
11.84517	12.23207	2.258448	-0.06397	0.018422	0.981747
11.31971	9.648881	1.820493	0.078965	0.03423	0.96635
11.13994	10.31276	1.923122	0.061641	0.079895	0.923214
10.88668	8.919196	2.127963	0.064265	0.058232	0.943431
11.21286	9.958359	1.9014	0.010486	0.001891	0.99811
10.32729	10.58073	1.750002	-0.06695	0.016298	0.983834
10.19991	8.780142	1.701385	0.055135	0.050476	0.950777
12.49303	11.00288	1.630248	0.161735	0.056437	0.945126
9.492011	10.18432	1.774155	-0.0021	0.14982	0.860863
12.35077	9.885959	2.06321	0.101466	0.012984	0.9871
13.83579	12.38751	2.210399	0.09591	0.023901	0.976383
10.7793	9.79506	1.914157	-0.002	0.022238	0.978007
13.5677	11.57484	2.056706	0.107292	0.002906	0.997099
7.985501	5.34217	2.037936	-0.02763	0.009152	0.990889
13.38003	11.66495	2.002266	0.128482	0.035195	0.965417
10.98503	10.92631	2.002894	-0.02832	0.042995	0.957916
10.29479	8.509924	2.136452	0.01426	0.041637	0.959218
12.01018	11.04497	1.824885	0.104352	0.068576	0.933722
11.4559	8.738427	2.172738	0.078226	0.016195	0.983935
10.1465	7.172985	1.881084	0.120919	0.045525	0.955496
11.0321	12.13576	1.901044	0.028695	0.135177	0.873562
9.997188	8.146266	2.094726	0.076457	0.106401	0.899064
8.942316	8.945452	2.01414	-0.10324	0.075202	0.927556
7.918214	5.939875	1.875064	-0.06426	0.013188	0.986898
13.05874	9.265323	2.12508	0.199189	0.029936	0.970508
9.316231	11.01349	1.611978	-0.112	0.07378	0.928876
11.81206	11.63631	2.599294	-0.03515	0.060221	0.941557
8.24342	7.741706	2.027592	-0.14936	0.039857	0.960927
13.56552	11.59484	2.076278	0.117617	0.016623	0.983515
10.24636	9.355291	1.964718	-0.00416	0.054334	0.947115
10.37671	11.14638	2.178873	-0.06753	0.091779	0.912307
14.06131	13.41037	2.331013	0.072463	0.037238	0.963447
11.71972	10.28001	1.772972	0.136335	0.08194	0.921328
7.077069	8.023909	1.942196	-0.17908	0.169835	0.843804
12.32743	9.893271	2.085019	0.087633	0.004041	0.995967
11.20332	8.690851	2.08392	0.092521	0.039621	0.961154
10.82558	11.25465	1.905639	-0.08147	0.006825	0.993198
9.850783	11.48707	2.433104	-0.16613	0.087836	0.915911
11.64799	10.69344	2.109627	-0.00708	0.007844	0.992187
12.98679	11.66243	1.793757	0.092682	0.001626	0.998375

Continued on next page

Table 1 – continued from previous page

Observed output y	Input x_1	Input x_2	v	Random noise u	True TE
9.129472	8.870868	1.895472	-0.1331	0.005258	0.994755
10.51448	13.94087	1.871405	-0.13526	0.084684	0.918802
12.49037	9.662434	2.18736	0.175528	0.078978	0.92406
9.689007	8.351853	1.984152	-0.0213	0.038856	0.961889
9.669802	8.913754	1.942321	0.028784	0.118157	0.888556
10.97188	11.70022	1.936571	-0.03523	0.063082	0.938866
10.87821	8.630425	2.040235	0.051063	0.018826	0.98135
13.79199	13.39889	2.159924	0.066692	0.031321	0.969165
11.60044	9.297126	1.87922	0.08018	0.000317	0.999683
10.82538	10.72119	1.612068	0.110076	0.132285	0.876091
13.7271	12.81615	2.267315	0.053944	0.013187	0.9869
10.45656	9.311159	2.266286	-0.06694	0.004584	0.995426
8.835307	11.58141	1.606668	-0.24529	0.017812	0.982346
8.779219	10.15401	1.94208	-0.1227	0.128397	0.879504
10.88107	9.025844	1.998324	0.067357	0.052068	0.949265
10.58934	13.19278	2.115319	-0.18053	0.035373	0.965246

Table 2: Estimates of technical efficiency scores under both models

TE for Normal-Truncated Skewed-Laplace	TE for Normal-Exponential
0.9513719	0.936188
0.9345194	0.997378
0.8873956	1.108503
0.9521405	0.932728
0.93549	0.99439
0.9557539	0.915326
0.8771736	1.1277
0.9581752	0.902525
0.8954012	1.092698
0.9573848	0.90681
0.9804154	0.704809
0.8375097	1.196364
0.9338301	0.999472
0.8853439	1.112435
0.9565324	0.911315
0.9672844	0.844302
0.9678795	0.839788
0.9453748	0.960733
0.9728307	0.797153
0.9539643	0.924187
0.962362	0.87797
0.9266853	1.019984
0.9595659	0.894725
0.9596518	0.894232

Continued on next page

Table 2 – continued from previous page

TE for Normal-Truncated Skewed-Laplace	TE for Normal-Exponential
0.9520285	0.933237
0.956242	0.912823
0.9592695	0.896416
0.8429304	1.187281
0.915257	1.049177
0.9301893	1.010181
0.9289003	1.013841
0.9361109	0.992454
0.9696816	0.825393
0.9464191	0.956738
0.9023461	1.078282
0.9711119	0.813101
0.9361648	0.992285
0.9509467	0.938068
0.9444146	0.964318
0.9530381	0.928584
0.9678452	0.840051
0.9531075	0.928259
0.8929104	1.0977
0.9660433	0.853378
0.9340346	0.998853
0.9633861	0.871442
0.9524135	0.93148
0.9571748	0.907931
0.9576263	0.905512
0.9362945	0.991878
0.9577642	0.904767
0.9702789	0.82036
0.9101242	1.061162
0.9683122	0.836433
0.9657558	0.855418
0.9514011	0.936058
0.9696359	0.825772
0.9510461	0.937631
0.9685352	0.83468
0.9388757	0.98358
0.9508644	0.938429
0.9617349	0.881859
0.9654295	0.857707
0.9677882	0.840487
0.9269666	1.019214
0.9505842	0.939652
0.8969241	1.089598
0.9409807	0.976517
0.9753541	0.770966

Continued on next page

Table 2 – continued from previous page

TE for Normal-Truncated Skewed-Laplace	TE for Normal-Exponential
0.8931101	1.097302
0.9290244	1.013491
0.8925709	1.098375
0.969272	0.828766
0.943401	0.968015
0.9037102	1.07536
0.9603325	0.890278
0.9645679	0.86362
0.7828696	1.28838
0.9677563	0.840732
0.9643906	0.864815
0.9337622	0.999677
0.8440264	1.185439
0.9525679	0.930769
0.9685324	0.834703
0.9168193	1.045405
0.8680057	1.144211
0.9690838	0.830295
0.9435089	0.967625
0.9347933	0.996539
0.9300205	1.010663
0.9616124	0.88261
0.9605924	0.888746
0.9677113	0.841075
0.952166	0.932612
0.9613915	0.883956
0.9390267	0.983083
0.8433233	1.186621
0.8508951	1.173831
0.9589107	0.898442
0.8697395	1.14113

2. The starting values of parameter are used to compute the true technical efficiency.

$s = 0.1$

$lambda = 0.03$

$p = 0.05$

$b0 = 3$

$b1 = 0.5$

$b2 = 0.25$

$x1 = rnorm(100, 10, 2)$

$x2 = rnorm(100, 2, 0.2)$

$v = rnorm(100, 0, s)$

3. The codes are used to estimate random noise u_i.

$t = runif(100, 0, 1)$

$mycdf = function(x, lambda, p, t)y = 1 + (exp(-(1 + lambda) * x/p)$

-2*(1+lambda)*exp(-x/p))/(2*lambda+1)-t

return(y)

$u = array(0, c(1, 100))$

$for(i\ in 1 : 100)$

$\{u[i] = uniroot(mycdf, c(0, 200), lambda = lambda, p = p, t = t[i])\$root$

4. The following codes are used to compute log likelihood function of normal-truncated skewed-Laplace model.

$likelihoodl = function(par)\{$

$s = par[1]$

$lambda = par[2]$

$p = par[3]$

$b0 = par[4]$

$b1 = par[5]$

$b2 = par[6]$

$e = log(data[,1]) - log(b0) - b1 * log(data[,2]) - b2 * log(data[,3])$

$s1 = (2 * (1 + lambda) * exp(e/p + s^2/2/p^2))/(p * (2 * lambda + 1)) * pnorm(-e/s - s/p)$

$s2 = ((1 + lambda) * exp(e * (1 + lambda)/p + ((1 + lambda)^2) * (s^2)/2/p^2))/(p * (2 * lambda + 1)) * pnorm(-e/s - (1 + lambda) * s/p)$

$logl = sum(log(s1 - s2))$

$return(logl)\}$

5. The following codes are used to find estimated values of each normal-truncated skewed-Laplace parameter. $s = seq(0.09, 0.13, 0.0025)$

$lambda = seq(0.02, 0.06, 0.0025)$

$p = seq(0.04, 0.08, 0.0025)$

$b0 = seq(2.9, 3.1, 0.01)$

$b1 = seq(0.45, 0.55, 0.01)$

$b2 = seq(0.2, 0.35, 0.01)$

$L = array(0, c(18158448, 1))$

$y = 0$

$for(i\ in 1 : 17)$

$for(j\ in 1 : 17)$

$for(k\ in 1 : 17)$

$for(l\ in 1:21)$

$for(m\ in 1:11)$

$for(n\ in 1:16)$

$\{y = y + 1$

$L[y] = likelihoodl(c(s[i], lambda[j], p[k], b0[l], b1[m], b2[n]))$

L

6. The codes to estimate technical efficiency of normal-truncated skewed-Laplace.

$su = 0.09$

$lambda = 0.06$

$sv = 0.065$

$epsilon = log(data[,1]) - log(2.91) - 0.52 * log(data[,2]) - 0.27 * log(data[,3])$

$s1 = sigma * exp(-epsilon^2/sigma^2/2) - 2 * (2 * pi)^(0.5) * exp(epsilon/p + sigma^2/p^2/2) * (epsilon + sigma^2/p) * pnorm(-epsilon/sigma - sigma/p) + (2 * pi)^(0.5) * exp(epsilon * (1 + lambda)/p + sigma^2 * (1 + lambda)^2/p^2/2) * (epsilon + (1 + lambda) * sigma^2/p) * pnorm(-epsilon/sigma - (1 + lambda) * sigma/p)$

$s2 = 2 * (2 * pi)^(0.5) * exp(epsilon/p + sigma^2/p^2/2) * pnorm(-epsilon/sigma - sigma/p) - (2 * pi)^(0.5) * exp(epsilon * (1 + lambda)/p + (1 + lambda) * sigma^2/p^2/2) * pnorm(-epsilon/sigma - (1 + lambda) * sigma/p)$

$TE = exp(-(s1/s2))$

7. The following codes are used to compute log likelihood function of normal-exponential model.

$likelihoodl\ nor\ exp = function(par)\{$

$su = par[1]$

$sv = par[2]$

$b0 = par[3]$

$b1 = par[4]$

$b2 = par[5]$

$e = log(data[,1]) - log(b0) - b1 * log(data[,2]) - b2 * log(data[,3])$

$logl = sum((e/su + sv^2/su^2/2) - log(su) + log(pnorm(-sv/su - e/sv)))$

$return(logl)\}$

8. The following codes are used to find estimated values of each normal-exponential parameter.

$su = seq(0.03, 0.07, 0.0025)$

$sv = seq(0.08, 0.12, 0.0025)$

$b0 = seq(2.9, 3.1, 0.01)$

$b1 = seq(0.45, 0.55, 0.01)$

$b2 = seq(0.2, 0.35, 0.01)$

$L = array(0, c(1068144, 1))$

$y = 0$

$for(i\ in 1 : 17)$

$for(k\ in 1 : 17)$

$for(l\ in 1:21)$

$for(m\ in 1:11)$

$for(n\ in 1:16)$

$\{y = y + 1$

$L[y] = likelihoodl\ nor\ exp(c(su[i], sv[k], b0[l], b1[m], b2[n]))\}$

L

9. The codes to estimate technical efficiency of normal-exponential.

$\sigma_u = 0.0675$

$\sigma_v = 0.09$

$e = log(data[, 1]) - log(2.91) - 0.52 * log(data[, 2]) - 0.27 * log(data[, 3])$

$s0 = sv * exp(-e^2/(sv^2/2) - e/su - sv^2/su^2)/(2 * pi)^(0.5) * pnorm(-e/sv - sv/su)$

$TE = exp(s0 - e - sv^2/su)$

Bibliography

[1] Afriat, S. N. Efficiency Estimation of Production Functions, *International Economic Review*, 13:3(October), 568-98, 1972.

[2] Aigner, D.J., C.A.K.Lovell, and P.Schmidt, Formulation and Estimation of Stochastic Frontier Production Function Models, *Journal of Econometrics*, 6:1(July), 21-37, 1977.

[3] Aigner, D.J., and S.F.Chu, On Estimation the Industry Production Function, *American Economic Review*, 58:4(September), 826-39, 1968.

[4] Arrow, K. J., H. B. Chenery, B. S. Minhas, and R. M. Solow Capital-Labor Substitution and Economic Efficiency, *Review of Economics and Statistics*, 63:3(August), 225-50, 1961.

[5] Bera, A. K., and S. C. Sharma, Estimating Production Uncertainty in Stochastic Frontier Production Frontier Models, Working Paper, Department of Economics, University of Illinois, Champaign, IL.

[6] Berndt, E. R., and L. R. Christensen, The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1928-1968, *Journal of Econometrics* 1:1(March), 81-114, 1973.

[7] Cobb, C., and P. H. Douglas, A Theory of Production, *American Economic Review Supplement*, 18, 139-65, 1928.

[8] Debrew, G., The Coefficient of Resource Utilization, *Econometrics*, 19:3(July), 272-92, 1951.

[9] Farrell, M. J., The Measurement of Productive Efficiency, *Journal of the Royal Statistical Society*, Series A, General, 120, Part 3, 253-81, 1957.

[10] Fried, H.O., C.A.K. Lovell, and S.S. Schmidt, eds. The Measurement of Production Efficiency: Techniques and Applications, *New York: Oxford University Press*, 1993.

[11] Gokarna Aryal, A.N.V.Rao, Reliability model using truncated skewed-Laplace distribution, *Nonlinear Analysis*, 63, 639-646, 2005.

[12] Greene, W.H., Maximum Likelihood Estimation of Econometric Frontier Functions, *Journal of Econometrics* , 13:1(May), 27-56, 1980(a).

[13] Greene, W.H., On the Estimation of a Flexible Frontier Production Model, *Journal of Econometrics* , 13:1(May), 101-15, 1980(b).

[14] Greene, W.H., A Gamma-Distributed Stochastic Frontier Model, *Journal of Econometrics*, 46:1/2(October/November), 141-64, 1990.

[15] Greene, W.H., Frontier Production Functions, in M.H.Pesaran and P. Schmidt, eds, *Handbook of Applied Econometrics, Volume II : Microeconomics*, Oxford: Blackwell Publishers, Ltd, 1997.

[16] Jondrow, J., C. A. K. Lovell, I. S. Materov, and P. Schmidt, On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model, *Journal of Econometrics*, 19:2/3, 233-38, 1982.

[17] Koopmans, T. C., An Analysis of Production as an Efficient Combination of Activities, in T. C. Koopmans, ed.,*Activity Analysis of Production and Allocation*, Cowles Commission for Research in Economics, Monograph No. 13, New York: Wiley.

[18] Luc Devroye, *Non-Uniform Random Variate Generation*, New York: Springer-Verlag, 1986.

[19] Meeusen, W., and J. van den Broeck, Efficiency Estimation from Cobb-Douglas Production Function with Composed Error, *International Econometric Review*, 18:2(June),435-44, 1977.

[20] Richmond, J., Estimating the Efficiency of Production, *The international Economic Review*, 15:2(June), 515-21, 1974.

[21] Schmidt, P., Frontier Production Functions, *Econometric Reviews*, 4:2, 289-328, 1985-1986.

[22] Stevenson, R. E., Likelihood Functions for Generalized Stochastic Frontier Estimation, *Journal of Econometrics*, 19:3(May), 57-66, 1980.

[23] Subal C. Kumbhakar, C.A. Knox Lovell, *Stochastic Frontier Analysis*, Cambridge University Press, 2003.

[24] William H. GREENE, On the estimation of a flexible frontier production model, *Journal of Econometrics*, 13, 101-115, 1980.