

A TEST FOR DISTRIBUTIONAL ASSUMPTIONS FOR THE  
STOCHASTIC FRONTIER FUNCTIONS

by

Lung-Fei Lee

Discussion Paper No. 81 - 151, May 1981

Center for Economic Research  
Department of Economics  
University of Minnesota  
Minneapolis, Minnesota 55455

# A TEST FOR DISTRIBUTIONAL ASSUMPTIONS FOR THE STOCHASTIC FRONTIER FUNCTIONS

By

Lung-Fei Lee\*

## 1. Introduction

Technically efficient production is defined as the maximum quantity of output attainable from given inputs. Knowledge of the production frontier, defined as the locus of technically efficient input-output combinations, and the actual input-output combinations of firms is sufficient information for measuring efficiency. A major difficulty is estimating the production frontier. Typically, empirical production functions are "average" rather than frontier functions, and thus unable to provide information on efficiency, because they attribute differences from the estimated function to the random disturbances. Attempts to estimate frontier production function began with the pioneering work of Farrell [1957] and subsequently, Aigner and Chu [1968], Afriat [1972], Richmond [1974] and others. They estimated the frontier using linear and quadratic programming techniques. Their approaches are based on deterministic frontiers. There are several disadvantages to their approaches. The most important problem is that it does not allow for random shocks in the production process which are outside the firm's control. As a consequence, a few extreme observations determine the frontier and exaggerate the maximum possible output

given inputs. Some problems occur for the estimation of other deterministic frontier functions such as the cost and profit frontier functions.

Recognizing this problem, Timmer [1971] eliminated a certain percentage of the total observations. Such a selection procedure, however, is not based on statistical theory and the number of observations eliminated is arbitrary. Recently, Aigner, Lovell and Schmidt [1977] and Meeusen and van der Broek [1977] handled this problem with a more satisfactory conceptual basis by explicitly including an error component, which is one-side distributed, in the overall disturbance, to capture the different inefficiency across the production units. They specified the distribution to be half-normal or exponential distributed and have applied the models to the Norwegian manufacturing in Meeusen et al. [1977] and the U.S. primary metals industry and U.S. agriculture in Aigner et al. [1977]. In the empirical applications, the fits of half-normal and exponential are of little difference. Subsequent applications of this stochastic frontier production function model adopt the half-normal distributions and are in Lee and Tyler [1978], Tyler and Lee [1979], Broeck et al. [1980], Pitt and Lee [1981] among others. While the empirical results are of interest, one can recognize that there may be a problem in the distributional assumption. Ideally, the specification of the efficient distributions should base on the information on the economic mechanisms generating the inefficiency. However, in the empirical applications, the econometricians do not have, in general, such information and they do not appear to have good a priori arguments to justify the choice of any particular one-sided distribution. Other one-side distributions can also be used. In Afriat [1972], Richmond [1974] and Greene [1980], gamma distribution is specified for the full (deterministic) frontier functions. In Stevenson [1980], the gamma distribution and truncated normal distribution

are specified for the stochastic frontier functions. The different specifications do give different estimates as demonstrated by the empirical examples in Richmond [1974], Broek et al. [1980], Greene [1980] and Stevenson [1980]. The differences are not only in the measure of efficiency but also on the estimates of the input coefficients in some cases. Given these ambiguous results, Broek et al. [1980, p. 137] and Forsund et al. [1980, p. 17] concluded that different approaches or different specified distributions lead to different results and they are not clear cut how the choice can be made.

Since there are no a priori arguments for the choice of a particular distribution, one needs to base the choice and evaluation by statistical means. In this article, we suggest the use of specification test to evaluate the specified one-side distribution. In the event that the test accepts the hypothesized distribution, one may have more confidence on the approach and the specification. If the data contradict sharply with the specific distribution, one should try some other flexible distributions. The approach we adopt is a Lagrangean multiplier (LM) test based on the truncated Pearson family. The Lagrangean multiplier test is of interest since it is computationally simple and is constructed from the estimated residuals. The truncated Pearson family of distributions is used since it contains many distributions of various shapes and the half normal, truncated normal, exponential and gamma distributions as special cases. In this article, we derive the tests for half normal and truncated normal distributions since these are the often used distributions for the estimation of stochastic frontier functions.

This article is organized as follows. In the second section, we specify the stochastic frontier function with the Pearson family of truncated distributions. In the third section, we derive a LM test for the half-normal

distribution and provide some interpretations of the test. In the fourth section, we extend our analysis to the truncated normal distribution. Finally, we draw our conclusions. An appendix is provided to verify the non-singularity of the information matrix used in the tests.

## 2. Stochastic Frontier Function and the Pearson Family of Truncated Distributions

The stochastic frontier function that we will consider is specified as

$$y_i = x_i \beta + \epsilon_i \quad i = 1, \dots, N \quad (2.1)$$

$$\epsilon_i = -u_i + v_i \quad u_i \geq 0$$

where  $y_i$  is the dependent variable;  $x_i$  is a  $k$  vector of exogenous variables which includes a constant term and the disturbances  $(u_i, v_i)$ ,  $i = 1, \dots, N$  are independently and identically distributed.<sup>1</sup> The disturbance  $u$  is one-side distributed and represents the inefficient component of the function while  $v$  represents the disturbance which cannot be controlled by the firm and is assumed to be statistically independent with the inefficient component. Since  $v_i$  represents uncontrollable random events and measurement error on the dependent variable, it is assumed to be normally distributed  $N(0, \sigma_v^2)$ . The important problem, however, is to specify an appropriate one-side distribution for  $u_i$ . Since  $u_i$  is the main interest in the model, we will focus our analysis on distributional assumption on  $u_i$ . In the econometric literature, half normal, exponential, truncated normal and gamma distributions have been specified for the inefficient component  $u$ . For our analysis, we assume that the distribution of  $u$  belongs to the Pearson family of truncated distributions.

The general density function of the Pearson family of truncated distributions can be expressed in the following form:

$$f(u) = \exp \left[ \frac{a+u}{b_0 + b_1 u + b_2 u^2} du \right] / \int_0^\infty \exp \left[ \frac{a+t}{b_0 + b_1 t + b_2 t^2} dt \right] dt \quad (2.2)$$

It is well known, see, e.g., Elderton and Johnson [1969] or Johnson and Kotz [1970], that the Pearson family of distributions contain distributions of various shapes and include the familiar normal, student t, beta, exponential and gamma distribution as special cases.<sup>2</sup> It follows immediately that the Pearson family of truncated distributions will contain the half-normal, truncated normal, exponential and gamma distribution as special cases. One can easily show that when  $b_1 = b_2 = a = 0$ , it is the half-normal distribution specified in Aigner et al. [1977]. When  $b_1 = b_2 = 0$ , it is the truncated normal distribution specified in Stevenson [1980]. The gamma distribution corresponds to the case that  $b_2 = 0$ . Since the family of distribution is so general and contains the distributions specified for the inefficient component, it provides a useful framework for testing the conventional distributional assumptions for the stochastic frontier functions. The uses of the Pearson family for other models can be found in Cohen [1959], Bera et al. [1980] and Lee [1981a, 1981b].

Under the assumptions that  $u$  belongs to the Pearson family of truncated distributions and  $v$  is normally distributed, the density function for the overall disturbance  $\varepsilon$  is

$$h(\varepsilon) = \int_0^\infty \frac{1}{\sigma_v} \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) \exp \left[ \frac{a+u}{b_0 + b_1 u + b_2 u^2} du \right] du / \int_0^\infty \exp \left[ \frac{a+t}{b_0 + b_1 t + b_2 t^2} dt \right] dt \quad (2.3)$$

where  $\phi(\cdot)$  is the standard normal density function. Let

$$p(t) \equiv \int \frac{a+t}{b_0 + b_1 t + b_2 t^2} dt$$

The log-likelihood function for the general model is

$$\begin{aligned} L &= \sum_{i=1}^N \ln h(y_i - x_i \beta) \\ &= \sum_{i=1}^N \left\{ \ln \left[ \int_0^{\infty} \frac{1}{\sigma_v} \phi \left( \frac{y_i - x_i \beta + u}{\sigma_v} \right) \exp(p(u)) du \right] - \ln \int_0^{\infty} \exp(p(t)) dt \right\} \end{aligned} \quad (2.4)$$

Let  $\varepsilon_i = y_i - x_i \beta$ . It is straightforward to show that the first order derivatives of  $L$  are

$$\frac{\partial L}{\partial \beta} = \sum_{i=1}^N \frac{x_i'}{\sigma_v^2} \left[ \varepsilon_i + \int_0^{\infty} u \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \middle/ \int_0^{\infty} \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \right] \quad (2.5)$$

$$\frac{\partial L}{\partial \sigma_v^2} = \sum_{i=1}^N \frac{1}{2\sigma_v^2} \int_0^{\infty} \left[ \frac{1}{\sigma_v} (\varepsilon_i + u)^2 - 1 \right] \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \middle/ \int_0^{\infty} \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \quad (2.6)$$

$$\begin{aligned} \frac{\partial L}{\partial a} &= \sum_{i=1}^N \left\{ \int_0^{\infty} \frac{\partial p(u)}{\partial a} \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \middle/ \int_0^{\infty} \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \right. \\ &\quad \left. - \int_0^{\infty} \frac{\partial p(u)}{\partial a} \exp(p(u)) du \middle/ \int_0^{\infty} \exp(p(u)) du \right\} \end{aligned} \quad (2.7)$$

$$\frac{\partial L}{\partial b_\ell} = \sum_{i=1}^N \left\{ \int_0^{\infty} \frac{\partial p(u)}{\partial b_\ell} \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \middle/ \int_0^{\infty} \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \exp(p(u)) du \right.$$

$$- \left. \int_0^{\infty} \frac{\partial p(u)}{\partial b_{\ell}} \exp(p(u)) du \right/ \int_0^{\infty} \exp(p(u)) du \quad \ell = 0,1,2 \quad (2.8)$$

where 
$$\frac{\partial p(u)}{\partial a} = \int \frac{1}{b_0 + b_1 u + b_2 u^2} du \quad (2.9)$$

and 
$$\frac{\partial p(u)}{\partial b_{\ell}} = - \int \frac{(a+u)u^{\ell}}{(b_0 + b_1 u + b_2 u^2)^2} du \quad \ell = 0,1,2 \quad (2.10)$$

These derivatives do not have closed form expressions since the solutions of the indefinite integrals involved depend on the characteristics of the solution of the quadratic equation  $b_0 + b_1 u + b_2 u^2 = 0$ . These derivatives provide the basis for the derivation of the LM test for any specific distribution in the family.

To justify the theoretical properties of the statistics, we assume that the parameter space is a compact set with the true parameter as an interior point, the empirical distribution of the vector of exogenous variables  $w_i$  in  $x_i = (1, w_i)$  converges to a limiting distribution  $F(w)$  and the variance-covariance matrix of  $w$  under  $F(w)$  is positive definite.

### 3. LM Test for Half-Normal Distribution in the Stochastic Frontier Function

If the distribution of the inefficient component  $u$  is assumed to be half normal  $N(0, \sigma_u^2)$ , the density function of  $u$  will be

$$f(u) = \frac{2}{\sqrt{2\pi} \sigma_u} \exp\left\{-\frac{u^2}{2\sigma_u^2}\right\} \quad u \geq 0 \quad (3.1)$$

This distribution corresponds to the case that  $a = 0$ ,  $b_1 = 0$  and  $b_2 = 0$  with  $b_0 = -\sigma_u^2$  in the Pearson family of truncated distributions. Under this

hypothesis  $H_N$ :  $a = 0$ ,  $b_1 = 0$  and  $b_2 = 0$ , the derivatives of the log likelihood function  $L$  in the equations (2.5)-(2.8) will be simplified as follows:

$$\left. \frac{\partial L}{\partial \beta} \right|_{H_N} = \sum_{i=1}^N \frac{x_i}{\sigma_v^2} \left\{ \varepsilon_i + \int_0^\infty u \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du \middle/ \int_0^\infty \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du \right\} \quad (3.2)$$

$$\left. \frac{\partial L}{\partial \sigma_v^2} \right|_{H_N} = \frac{1}{2\sigma_v^2} \sum_{i=1}^N \left\{ \frac{1}{\sigma_v} \frac{\int_0^\infty (\varepsilon_i + u)^2 \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du}{\int_0^\infty \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du} - 1 \right\} \quad (3.3)$$

$$\left. \frac{\partial L}{\partial a} \right|_{H_N} = -\frac{1}{\sigma_u^2} \sum_{i=1}^N \left\{ \frac{\int_0^\infty u \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du}{\int_0^\infty \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du} - \frac{\int_0^\infty u \phi \left( \frac{u}{\sigma_u} \right) du}{\int_0^\infty \phi \left( \frac{u}{\sigma_u} \right) du} \right\} \quad (3.4)$$

$$\left. \frac{\partial L}{\partial b_0} \right|_{H_N} = -\frac{1}{(2+l)\sigma_u^4} \sum_{i=1}^N \left\{ \frac{\int_0^\infty u^{2+l} \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du}{\int_0^\infty \phi \left( \frac{\varepsilon_i + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du} - \frac{\int_0^\infty u^{2+l} \phi \left( \frac{u}{\sigma_u} \right) du}{\int_0^\infty \phi \left( \frac{u}{\sigma_u} \right) du} \right\} \quad (3.5)$$

$l = 0, 1, 2$

Let

$$g_1(\varepsilon) = (2\pi)^{-1/2} (\sigma_u^2 + \sigma_v^2)^{-1/2} \exp \left\{ -\frac{\varepsilon^2}{2(\sigma_u^2 + \sigma_v^2)} \right\} \quad (3.6)$$

be the density function of a normal random variable  $N(0, \sigma_u^2 + \sigma_v^2)$  and

$$g_2(u|\varepsilon) = (2\pi)^{-1/2} (\sigma_u \sigma_v)^{-1} (\sigma_u^2 + \sigma_v^2)^{1/2} \exp \left[ - \frac{\sigma_u^2 + \sigma_v^2}{2\sigma_u^2 \sigma_v^2} \left( u + \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} \varepsilon \right)^2 \right] \quad (3.7)$$

be the normal density function of another random variable

$$N \left[ - \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} \varepsilon, \frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2} \right]$$

One can easily show that

$$\phi \left( \frac{\varepsilon + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) = \sigma_u \sigma_v g_1(\varepsilon) g_2(u|\varepsilon) \quad (3.8)$$

Let

$$g^*(u|\varepsilon) = g_2(u|\varepsilon) / \int_0^\infty g_2(u|\varepsilon) du \quad u \geq 0 \quad (3.9)$$

be the truncated normal density function defined on the non-negative values corresponding to the normal density function of  $g_2(u|\varepsilon)$ . It follows immediately from (3.8) and (3.9) that

$$\phi \left( \frac{\varepsilon + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) / \int_0^\infty \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) \phi \left( \frac{u}{\sigma_u} \right) du = g^*(u|\varepsilon) \quad (3.10)$$

Also

$$\phi \left( \frac{u}{\sigma_u} \right) / \int_0^\infty \phi \left( \frac{u}{\sigma_u} \right) du = f(u) \quad (3.11)$$

Hence the derivatives in (3.2)-(3.5) are equivalent to

$$\left. \frac{\partial L}{\partial \beta} \right|_{H_N} = \sum_{i=1}^N \frac{x_i'}{\sigma_v^2} \left\{ \varepsilon_i + \int_0^\infty u g^*(u|\varepsilon_i) du \right\} \quad (3.2)'$$

$$\left. \frac{\partial L}{\partial \sigma_v^2} \right|_{H_N} = \frac{1}{2\sigma_v^3} \sum_{i=1}^N \left\{ \varepsilon_i^2 + 2\varepsilon_i \int_0^\infty u g^*(u|\varepsilon_i) du + \int_0^\infty u^2 g^*(u|\varepsilon_i) du - \sigma_v \right\} \quad (3.3)'$$

$$\left. \frac{\partial L}{\partial a} \right|_{H_N} = -\frac{1}{\sigma_u^2} \sum_{i=1}^N \left\{ \int_0^\infty u g^*(u|\varepsilon_i) du - \int_0^\infty u f(u) du \right\} \quad (3.4)'$$

$$\left. \frac{\partial L}{\partial b_\lambda} \right|_{H_N} = -\frac{1}{(2+\lambda)\sigma_u^4} \sum_{i=1}^N \left\{ \int_0^\infty u^{2+\lambda} g^*(u|\varepsilon_i) du - \int_0^\infty u^{2+\lambda} f(u) du \right\} \quad (3.5)'$$

$\lambda = 0, 1, 2$

To derive a test for the hypothesis  $H_N$ , we suggest the use of the Lagrangean Multiplier (LM) test. The Lagrangean Multiplier test as a principle of testing hypothesis is proposed in Aitchison and Silvey [1958] and Silvey [1959] which generalizes the efficient score test of Rao [1973] and is related to the Neyman's  $C(\alpha)$  test [1959]. It is well known that the LM test has optimal asymptotic properties in the sense of Neyman (see Neyman [1959]) and is asymptotically equivalent to the maximum likelihood ratio test under general regularity conditions (see, e.g., Silvey [1959]). The usefulness of the LM test in econometrics has been demonstrated in Breusch and Pagan [1980], among others. Let  $\theta = (\beta', \sigma_v^2, a, b_0, b_1, b_2)'$  be the vector of parameters and

$$I(\theta) = E_{H_N} \left[ -\frac{\partial^2 L}{\partial \theta \partial \theta'} \right]$$

be the information matrix under the null hypothesis. The LM statistic for testing  $H_N$  is the statistic

$$\frac{\partial L}{\partial \theta'} (I(\theta))^{-1} \left. \frac{\partial L}{\partial \theta} \right|_{\theta=\hat{\theta}} \quad (3.12)$$

evaluated at the restricted MLE  $\hat{\theta}$ , which is asymptotically chi-square distributed with three degrees of freedom und  $H_N$ . If  $I(\theta)$  were singular, the inverse of  $I(\hat{\theta})$  should be replaced by the generalized inverse of  $I(\hat{\theta})$  and the degree of freedom would be reduced by the nullity of the matrix  $I(\theta)$ . It is known that the asymptotic properties of the statistics will not be affected if the negative of the Hessian matrix or its asymptotic equivalent form are used instead of the information matrix in constructing the test statistics.

For our model, it is difficult to derive the analytic information matrix under the hypothesis  $H_N$ , since  $\varepsilon_i$  appears in the derivatives in a highly nonlinear form. As an alternative, we suggest an estimate of the information matrix derived from the first derivatives of the log likelihood function. Let

$$\begin{aligned} l_{\beta i} &= x_i' \left[ \varepsilon_i + \int_0^{\infty} u g^*(u | \varepsilon_i) du \right] \\ l_{\sigma i} &= \varepsilon_i^2 + 2\varepsilon_i \int_0^{\infty} u g^*(u | \varepsilon_i) du + \int_0^{\infty} u^2 g^*(u | \varepsilon_i) du - \sigma_v \\ l_{a i} &= \int_0^{\infty} u g^*(u | \varepsilon_i) du - \int_0^{\infty} u f(u) du \\ l_{b_j i} &= \int_0^{\infty} u^{2+j} g^*(u | \varepsilon_i) du - \int_0^{\infty} u^{2+j} f(u) du, \quad j = 0, 1, 2 \end{aligned} \tag{3.13}$$

and define a diagonal matrix  $A$  of dimension  $k + 5$  as

$$A = \text{Diag} \left[ \sigma_v^{-2}, \dots, \sigma_v^{-2}, \frac{1}{2} \sigma_v^{-3}, -\sigma_u^{-2}, -\frac{1}{2} \sigma_u^{-4}, -\frac{1}{3} \sigma_u^{-4}, -\frac{1}{4} \sigma_u^{-4} \right] \tag{3.14}$$

Denote  $l_{\theta i} = (l'_{\beta i}, l_{\sigma i}, l_{a i}, l_{b_0 i}, l_{b_1 i}, l_{b_2 i})'$ . An estimate of the information matrix evaluated at  $\hat{\theta}$  can be constructed as

$$\Omega = A \left[ \sum_{i=1}^N \ell_{\theta i} \ell_{\theta i}' \right] A \quad (3.15)$$

evaluated at  $\hat{\theta}$ . Thus an asymptotically equivalent LM statistic for testing  $H_N$  is

$$\frac{\partial L}{\partial \theta'} \Omega^{-1} \frac{\partial L}{\partial \theta} \Big|_{\theta=\hat{\theta}} \quad (3.16)$$

In the appendix, we have given a proof that the limiting information matrix of  $\frac{1}{N}I(\theta)$  under  $H_N$  as  $N$  goes to infinity is nonsingular with probability one. So the LM statistic (3.16) is asymptotically chi-square distributed with three degrees of freedom under the null hypothesis  $H_N$ .

It is of interest to note that one can provide a simple interpretation of this statistic. The first-order derivatives in (3.2)'-(3.5)' can be rewritten in terms of  $\ell_{\theta i}$  and  $A$ ,

$$\frac{\partial L}{\partial \theta} \Big|_{H_N} = A \sum_{i=1}^N \ell_{\theta i}$$

The LM statistic as constructed in (3.16) can be simplified to

$$\left[ \sum_{i=1}^N \hat{\ell}_{\theta i} \right]' \left[ \sum_{i=1}^N \hat{\ell}_{\theta i} \hat{\ell}_{\theta i}' \right]^{-1} \left[ \sum_{i=1}^N \hat{\ell}_{\theta i} \right] \quad (3.17)$$

where  $\hat{\ell}_{\theta i}$  is the vector  $\ell_{\theta i}$  evaluated at  $\theta = \hat{\theta}$ . Since  $\hat{\theta}$  is the restricted MLE of  $\theta$  under  $H_N$ , the derivatives

$$\frac{\partial L}{\partial \beta} \Big|_{\theta=\hat{\theta}}, \quad \frac{\partial L}{\partial \sigma_V^2} \Big|_{\theta=\hat{\theta}}, \quad \text{and} \quad \frac{\partial L}{\partial b_0} \Big|_{\theta=\hat{\theta}}$$

evaluated at  $\hat{\theta}$  are zero. These imply that  $\sum_{i=1}^N \hat{\ell}_{\beta i}$  and  $\sum_{i=1}^N \hat{\ell}_{\sigma i}$  and  $\sum_{i=1}^N \hat{\ell}_{b_0 i}$  are also zero. Thus the LM statistic becomes

$$\left( \sum_{i=1}^N \hat{x}_{ai}, \sum_{i=1}^N \hat{x}_{b1i}, \sum_{i=1}^N \hat{x}_{b2i} \right) \left[ L \left( \sum_{i=1}^N \hat{x}_{\theta i} \hat{x}'_{\theta i} \right)^{-1} L' \right] \left( \sum_{i=1}^N \hat{x}_{ai}, \sum_{i=1}^N \hat{x}_{b1i}, \sum_{i=1}^N \hat{x}_{b2i} \right)' \quad (3.18)$$

where  $L$  is a  $3 \times (k+5)$  selection matrix which deletes the columns and rows corresponding to  $\sum_{i=1}^N \hat{x}_{\beta i}$ ,  $\sum_{i=1}^N \hat{x}_{\sigma i}$  and  $\sum_{i=1}^N \hat{x}_{b0i}$  from the matrix  $\left( \sum_{i=1}^N \hat{x}_{\theta i} \hat{x}'_{\theta i} \right)^{-1}$ .

Simple interpretations can be provided for the components:

$$\frac{1}{N} \sum_{i=1}^N x_{ai} = \frac{1}{N} \sum_{i=1}^N \int_0^{\infty} u g^*(u | \varepsilon_i) du - \int_0^{\infty} u f(u) du \quad (3.19)$$

$$\frac{1}{N} \sum_{i=1}^N x_{b1i} = \frac{1}{N} \sum_{i=1}^N \int_0^{\infty} u^3 g^*(u | \varepsilon_i) du - \int_0^{\infty} u^3 f(u) du \quad (3.20)$$

and

$$\frac{1}{N} \sum_{i=1}^N x_{b2i} = \frac{1}{N} \sum_{i=1}^N \int_0^{\infty} u^4 g^*(u | \varepsilon_i) du - \int_0^{\infty} u^4 f(u) du \quad (3.21)$$

Under the null hypothesis  $H_N$ , the joint density of  $\varepsilon$  and  $u$  is

$$\frac{1}{\sigma_v} \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) f(u)$$

and the marginal density is

$$\int_0^{\infty} \frac{1}{\sigma_v} \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) f(u) du$$

Hence the conditional density of  $u$  conditional on  $\varepsilon$  under  $H_N$  is

$$\frac{1}{\sigma_v} \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) f(u) \Bigg/ \int_0^{\infty} \frac{1}{\sigma_v} \phi \left( \frac{\varepsilon + u}{\sigma_v} \right) f(u) du$$

which is exactly the function  $g^*(u | \varepsilon)$  as shown in (3.10). Hence the truncated normal density function of  $g^*(u | \varepsilon)$  is in fact the conditional density function

of  $u$  given  $\varepsilon$  under  $H_N$ . Since  $f(u)$  is the hypothesized distribution of  $u$  under  $H_N$ , given samples  $\{\varepsilon_i\}$  of  $\varepsilon$ , the components in (3.19), (3.20) and (3.21) are the differences between the first, third and fourth moments of the conditional density function of  $u$  and the corresponding first, third and fourth moments of the unconditional density function of  $u$  under  $H_N$ . The LM statistic utilizes the estimated restricted MLE residuals  $\hat{\varepsilon}_i = y_i - x_i \hat{\beta}$  and the estimated differences between the first, third and fourth sample moments constructed from the conditional density of  $u$  and the corresponding first, third and fourth moments constructed from the hypothesized density function of  $u$ .<sup>3</sup>

The computation of the LM statistic involves the computation of the moments of the truncated normal density function  $g^*(u|\varepsilon)$  and the half-normal density function  $f(u)$ . The computations of these moments, however, are quite simple as the following recursive formulae can be used:

$$\int_0^{\infty} u f(u) du = \sigma_u^2 f(0) = 2\sigma_u \phi(0) \quad (3.22)$$

$$\int_0^{\infty} u^{n+1} f(u) du = n\sigma_u^2 \int_0^{\infty} u^{n-1} f(u) du \quad n \geq 1 \quad (3.23)$$

and

$$\int_0^{\infty} u g^*(u|\varepsilon) du = \frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2} \frac{g_2(0|\varepsilon)}{1 - \Phi\left(\frac{\sigma_u \varepsilon}{\sigma_v (\sigma_u^2 + \sigma_v^2)^{1/2}}\right)} - \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} \varepsilon \quad (3.24)$$

$$\int_0^{\infty} u^{n+1} g^*(u|\varepsilon) du = n \frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2} \int_0^{\infty} u^{n-1} g^*(u|\varepsilon) du - \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2} \varepsilon \int_0^{\infty} u^n g^*(u|\varepsilon) du, \quad n \geq 1 \quad (3.25)$$

where  $\phi(\cdot)$  is the standard normal density function. The above recursive formulae can be easily derived from the differential equation method. The details can be found in Cohen [1951] and Lee [1981].

#### 4. LM Test for Truncated Normal Distribution in the Stochastic Frontier Function

The half-normal distribution for the inefficient component  $u$  restricts the mode of the distribution to occur at  $u = 0$ . Stevenson [1980] suggests the use of truncated normal distribution for  $u$  so that the mode need not necessarily be zero. He has also pointed out the possibility of using gamma distribution for  $u$ . However, since only the truncated normal distribution has been used in practice, we will consider the LM test for this distribution only. The density function of the truncated normal random variable is

$$f(u) = \frac{1}{\sqrt{2\pi} \sigma_u} \exp\left\{-\frac{1}{2} \left(\frac{u-\mu}{\sigma_u}\right)^2\right\} / \Phi(\mu/\sigma_u), \quad u \geq 0 \quad (4.1)$$

where  $\phi(\cdot)$  is the standard normal distribution function. This distribution corresponds to the case that  $b_1 = 0$ ,  $b_2 = 0$ ,  $b_0 = \sigma_u^2$ ,  $a = -\mu$  in the Pearson family of truncated distribution. Hence a test for truncated normality is to test the hypothesis  $H_T$ :  $b_1 = 0$  and  $b_2 = 0$ . Define

$$h(u|\varepsilon) = (2\pi)^{-1/2} (\sigma_u \sigma_v)^{-1} (\sigma_u^2 + \sigma_v^2)^{1/2} \exp\left[-\frac{\sigma_u^2 + \sigma_v^2}{2\sigma_u^2 \sigma_v^2} \left(u + \frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2} \left(\frac{\varepsilon}{\sigma_v^2} + \frac{a}{\sigma_u^2}\right)\right)^2\right] \quad (4.2)$$

and

$$h^*(u|\varepsilon) = h(u|\varepsilon) / \int_0^\infty h(u|\varepsilon) du \quad (4.3)$$

It can be shown that  $h^*(u|\varepsilon)$  is the conditional density function of  $u$  conditional on  $\varepsilon$  when the hypothesis  $H_T$  is true. Under the hypothesis  $H_T$ , the

first order derivatives will be simplified to

$$\left. \frac{\partial L}{\partial \beta} \right|_{H_T} = \frac{1}{\sigma_v^2} \sum_{i=1}^N x_i' \left\{ \varepsilon_i + \int_0^\infty u h^*(u|\varepsilon_i) du \right\} \quad (4.4)$$

$$\left. \frac{\partial L}{\partial \sigma^2} \right|_{H_T} = \frac{1}{2\sigma_v^3} \sum_{i=1}^N \left\{ \varepsilon_i^2 + 2\varepsilon_i \int_0^\infty u h^*(u|\varepsilon_i) du + \int_0^\infty u^2 h^*(u|\varepsilon_i) du - \sigma_v \right\} \quad (4.5)$$

$$\left. \frac{\partial L}{\partial a} \right|_{H_T} = -\frac{1}{\sigma_u^2} \sum_{i=1}^N \left\{ \int_0^\infty u h^*(u|\varepsilon_i) du - \int_0^\infty u f(u) du \right\} \quad (4.6)$$

$$\begin{aligned} \left. \frac{\partial L}{\partial b_\ell} \right|_{H_T} = & -\frac{1}{\sigma_u} \sum_{i=1}^N \left\{ \frac{1}{2+\ell} \left[ \int_0^\infty u^{2+\ell} h^*(u|\varepsilon_i) du - \int_0^\infty u^{2+\ell} f(u) du \right] \right. \\ & \left. + \frac{a}{1+\ell} \left[ \int_0^\infty u^{1+\ell} h^*(u|\varepsilon_i) du - \int_0^\infty u^{1+\ell} f(u) du \right] \right\} \\ & \ell = 0, 1, 2 \quad (4.7) \end{aligned}$$

Let

$$l_{\beta i} = x_i' \left\{ \varepsilon_i + \int_0^\infty u h^*(u|\varepsilon_i) du \right\} \quad (4.8)$$

$$l_{\sigma i} = \varepsilon_i^2 + 2\varepsilon_i \int_0^\infty u h^*(u|\varepsilon_i) du + \int_0^\infty u^2 h^*(u|\varepsilon_i) du - \sigma_v \quad (4.9)$$

$$l_{a i} = \int_0^\infty u h^*(u|\varepsilon_i) du - \int_0^\infty u f(u) du \quad (4.10)$$

$$l_{b_\ell} = \int_0^\infty u^{2+\ell} h^*(u|\varepsilon_i) du - \int_0^\infty u^{2+\ell} f(u) du, \quad \ell = 0, 1, 2 \quad (4.11)$$

and  $l_{\theta i} = (l_{\beta i}', l_{\sigma i}, l_{a i}, l_{b_0 i}, l_{b_1 i}, l_{b_2 i})'$ . Let  $\tilde{\theta}$  be the restricted MLE of  $\theta$  under  $H_T$  and  $\hat{l}_{\theta i}$  be the vector  $l_{\theta i}$  evaluated at  $\theta = \tilde{\theta}$ . Since  $\tilde{\theta}$  is the

restricted MLE of  $\theta$ , the derivatives

$$\left. \frac{\partial L}{\partial \beta} \right|_{\theta=\tilde{\theta}}, \left. \frac{\partial L}{\partial \sigma^2} \right|_{\theta=\tilde{\theta}}, \left. \frac{\partial L}{\partial a} \right|_{\theta=\tilde{\theta}} \text{ and } \left. \frac{\partial L}{\partial b_0} \right|_{\theta=\tilde{\theta}}$$

evaluated at  $\tilde{\theta}$  are zero. These imply that the terms  $\sum_{i=1}^N \hat{\ell}_{\beta i}$ ,  $\sum_{i=1}^N \hat{\ell}_{\sigma i}$ ,  $\sum_{i=1}^N \hat{\ell}_{a i}$ , and  $\sum_{i=1}^N \hat{\ell}_{b_0 i}$  are all zero. By similar arguments in the previous section, the LM statistic for testing the hypothesis  $H_T$  can be constructed as

$$\left( \sum_{i=1}^N \hat{\ell}_{b_1 i}, \sum_{i=1}^N \hat{\ell}_{b_2 i} \right) \left[ L_1 \left( \sum_{i=1}^N \hat{\ell}_{\theta i} \hat{\ell}'_{\theta i} \right)^{-1} L_1' \right] \left( \sum_{i=1}^N \hat{\ell}_{b_1 i}, \sum_{i=1}^N \hat{\ell}_{b_2 i} \right)' \quad (4.12)$$

where  $L_1$  is a  $2 \times (h + 5)$  selection matrix which selects the last two rows of  $\left( \sum_{i=1}^N \hat{\ell}_{\theta i} \hat{\ell}'_{\theta i} \right)^{-1}$ . As compared with the test of half-normal distribution, the LM test for truncated normal distribution utilizes only the third and fourth moments of the hypothesized truncated normal distribution of  $u$  and the conditional distribution of  $u$  conditional on  $\varepsilon$ . The estimated differences between the third and fourth moments of the conditional distribution given the samples and the corresponding moments of the specified unconditional distribution form the basis of the LM test. The LM statistic is asymptotically chi-square distributed with two degrees of freedom.

The moments in the equations (4.8)-(4.11) can be computed by the following recursive formulae:

$$\int_0^{\infty} u f(u) du = -a + \sigma_u \phi(a/\sigma_u) / \Phi(-a/\sigma_u) \quad (4.13)$$

$$\int_0^{\infty} u^{n+1} f(u) du = n\sigma_u^2 \int_0^{\infty} u^{n-1} f(u) du - a \int_0^{\infty} u^n f(u) du, \quad n \geq 1 \quad (4.14)$$

and

$$\int_0^{\infty} u h^*(u|\varepsilon) du = \frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2} \left[ \frac{h(0|\varepsilon)}{1 - \Phi \left[ \frac{\sigma_u \sigma_v}{(\sigma_u^2 + \sigma_v^2)^{1/2}} \left( \frac{\varepsilon}{\sigma_v^2} + \frac{a}{\sigma_u^2} \right) \right]} - \frac{\varepsilon}{\sigma_v^2} - \frac{a}{\sigma_u^2} \right] \quad (4.15)$$

$$\int_0^{\infty} u^{n+1} h^*(u|\varepsilon) du = \frac{\sigma_u^2 \sigma_v^2}{\sigma_u^2 + \sigma_v^2} \left[ n \int_0^{\infty} u^{n-1} h^*(u|\varepsilon) du - \left( \frac{\varepsilon}{\sigma_v^2} + \frac{a}{\sigma_u^2} \right) \int_0^{\infty} u^n h^*(u|\varepsilon) du \right] \quad (4.16)$$

$n \geq 1$

## 5. Conclusion

In this article, we have derived some tests for testing distributional assumptions for the stochastic frontier functions. Since there are, in general, no a priori reasons to choose any particular one-side distribution for the inefficient component in the stochastic frontier functions and the empirical results are different for different distributional assumptions, it is important to rely on statistical means to evaluate the assumptions. The testing procedures we proposed are Lagrangean Multiplier tests. We suggest the use of the Pearson family of truncated distributions as the basis for our tests since it contains a broad class of distributions of different shapes and in particular the half-normal, truncated normal, exponential and gamma distributions as special cases. We have explicitly derived the tests for the half-normal and truncated normal distributions. The statistic utilizes the differences between the first, third and fourth moments of the conditional distribution of the inefficient component conditional on the samples and the corresponding moments of the hypothesized distribution for the half-normal distribution case. It

utilizes only the corresponding third and fourth moments for the testing of the truncated normal distribution.<sup>4</sup> The statistics are asymptotically chi-square distributed and can be easily computed. From our analysis, it is expected that the approach can be generalized to the testing of any specific distribution in the Pearson family.

Appendix  
Non-Singularity of the Information Matrix

The degree of freedom of the LM statistic depends on the rank of the information matrix under the null hypothesis. In this appendix, we will prove that the limiting matrix of  $\frac{1}{N}I(\theta)$  under the hypothesis  $H_N$  is of full rank. Similar arguments can be applied to the case under the hypothesis  $H_T$ .

Since analytic expression for  $I(\theta)$  is not available, we cannot analyze directly this matrix. However, one can analyze the following vector of random variables:

$$z_1 = \frac{x}{\sigma_v^2} \left[ \varepsilon + \int_0^\infty u g^*(u|\varepsilon) du \right] \quad (\text{A.1})$$

$$z_2 = \frac{1}{2\sigma_v^3} \left[ \varepsilon + 2\varepsilon \int_0^\infty u g^*(u|\varepsilon) du + \int_0^\infty u^2 g^*(u|\varepsilon) du - \sigma_v \right] \quad (\text{A.2})$$

$$z_3 = -\frac{1}{\sigma_u^2} \left[ \int_0^\infty u g^*(u|\varepsilon) du - \int_0^\infty u f(u) du \right] \quad (\text{A.3})$$

$$z_{4+\ell} = -\frac{1}{(2+\ell)\sigma_u^4} \left[ \int_0^\infty u^{2+\ell} g^*(u|\varepsilon) du - \int_0^\infty u^{2+\ell} f(u) du \right],$$

$$\ell = 0, 1, 2 \quad (\text{A.4})$$

which are the basic elements in the derivatives in (3.2)'-(3.5)'. Since the limiting matrix of  $\frac{1}{N}I(\theta)$  is the variance-covariance matrix of the random variables  $z_1, z_2, \dots, z_6$ , it is enough to show that the random variables  $z_1, z_2, \dots, z_6$  are not linearly dependent. This is so, since it is known that the variance-covariance matrix of a vector of random variables is positive

definite if and only if there is no linear relation among the random variables with probability one (see, e.g., Rao [1973], p. 107). Let  $c = (c_1, c_2, \dots, c_6)'$  be a vector of constants such that  $\sum_{i=1}^6 c_i z_i = 0$ . We would like to show that  $c$  is a zero vector.

Let

$$\lambda = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_v^2}$$

From the recursive relations in the equations (3.24) and (3.25), we have

$$\int_0^\infty u g^*(u|\epsilon) du = \lambda \sigma_v^2 g^*(0|\epsilon) - \lambda \epsilon \quad (\text{A.5})$$

$$\int_0^\infty u^2 g^*(u|\epsilon) du = \lambda^2 \epsilon^2 - \sigma_v^2 \lambda^2 g^*(0|\epsilon) \epsilon + \lambda \sigma_v^2 \quad (\text{A.6})$$

$$\int_0^\infty u^3 g^*(u|\epsilon) du = -\lambda^3 \epsilon^3 + \sigma_v^2 \lambda^3 g^*(0|\epsilon) \epsilon^2 - 3\sigma_v^2 \lambda^2 \epsilon + 2\sigma_v^4 \lambda^2 g^*(0|\epsilon) \quad (\text{A.7})$$

and

$$\begin{aligned} \int_0^\infty u^4 g^*(u|\epsilon) du = \lambda^4 \epsilon^4 - \sigma_v^2 \lambda^4 g^*(0|\epsilon) \epsilon^3 + 6\sigma_v^2 \lambda^3 \epsilon^2 - (2 + 3\sigma_v^2 \lambda) \sigma_v^2 \lambda^2 g^*(0|\epsilon) \epsilon \\ + 3\sigma_v^4 \lambda^2 \quad (\text{A.8}) \end{aligned}$$

Since the polynomial term  $\epsilon^4$  appears only in  $z_6$ ,  $c_6$  must be zero. It follows that  $\sum_{i=1}^5 c_i z_i = 0$  and  $c_5$  must be zero since the function  $\epsilon^3$  appears only in  $z_5$ . So the remained equality is

$$\sum_{i=1}^4 c_i z_i = 0 \quad (\text{A.9})$$

This equation (A.9) is a linear combination of functions of  $\epsilon$ . The equality in

(A.9) can hold for all possible values of  $\varepsilon$  under  $H_N$  if and only if all the coefficients of the functions of  $\varepsilon$ , namely  $\varepsilon^2$ ,  $g^*(o|\varepsilon)\varepsilon$ ,  $g^*(o|\varepsilon)$  and  $\varepsilon$ , respectively, vanish, i.e.,

$$-\frac{1}{2\sigma_u^4} \lambda^2 c_4 + \left( \frac{1}{2} - \lambda + \frac{1}{2} \lambda^2 \right) \frac{1}{\sigma_v^3} c_2 = 0 \quad (\text{A.10})$$

$$\frac{\sigma_v^2}{2\sigma_u^4} \lambda^2 c_4 + \frac{\lambda}{\sigma_v} \left( 1 - \frac{1}{2} \lambda \right) c_2 = 0 \quad (\text{A.11})$$

$$\frac{\sigma_v^2}{\sigma_u^2} \lambda c_3 - \lambda c_1' x = 0 \quad (\text{A.12})$$

$$\frac{\lambda}{\sigma_u^2} c_3 + \frac{1}{\sigma_v^2} (1 - \lambda) c_1' x = 0 \quad (\text{A.13})$$

Since

$$\left( \frac{1}{2} - \lambda + \frac{1}{2} \lambda^2 \right) / \sigma_v^3 = \frac{\sigma_v}{2(\sigma_u^2 + \sigma_v^2)^2}$$

the equation (A.10) is equivalent to

$$-c_4 + \sigma_v c_2 = 0 \quad (\text{A.10})'$$

The equation (A.11) is equivalent to

$$\sigma_v^3 c_4 + (\sigma_u^2 + 2\sigma_v^2) \sigma_u^2 c_2 = 0 \quad (\text{A.11})'$$

Obviously, the solutions of  $c_2$  and  $c_4$  from (A.10)' and (A.11)' must be zero.

Hence it remains to show that  $c_3$  and  $c_1$  are zero. The equation (A.12) is equivalent to

$$\sigma_v^2 c_3 - \sigma_u^2 c_1' x = 0 \quad (\text{A.12})'$$

and the equation (A.13) is equivalent to

$$c_3 + c_1' x = 0 \tag{A.13}'$$

They imply that  $(\sigma_u^2 + \sigma_v^2) c_3 = 0$  and hence  $c_3$  must be zero. It follows  $c_1' x = 0$ . Since  $x = (1, w')$  and  $w$  has a positive definite variance-covariance matrix under the distribution  $F(w)$  by our assumption, the vector  $c_1$  must be zero. Therefore,  $c = 0$  and the random variables  $z_1, \dots, z_6$  are not linearly dependent.

Footnotes

- \* The author is an associate professor of economics, University of Minnesota, and a visiting associate professor for the Center of Econometrics and Decision Sciences, University of Florida. Support from the National Science Foundation under grant SES 8006481 to the University of Minnesota is gratefully acknowledged. A preliminary version of this paper has been presented at the conference on current issues in productivity at Rutgers University, New Jersey, April 20-24. I appreciate having valuable comments from Professor A. Dogramaci. Any errors are of my own.
1. This specification with  $u \geq 0$  is applicable to the production and profit functions. The cost function  $y = x\beta + u + v$  with  $u \geq 0$  can be transformed to this function (2.1) by multiplying both sides by minus 1.
  2. Alternatively, the Pearson family of distributions is characterized by the differential equation  $d(\ln f(u))/du = (a + u)/(b_0 + b_1u + b_2u^2)$ . The solutions of  $f(u)$  depend on the characteristics of the solutions of the quadratic form  $b_0 + b_1u + b_2u^2$ . The family of distributions is divided into three main types and ten transition types.
  3. This quantity  $\int_0^\infty u g^*(u|\epsilon_i) du$  may be used as a predictor of the inefficiency level  $u_i$  for the production unit  $i$  for the stochastic frontier function. This observation is due to Professor Peter Schmidt and was pointed out to me in our conversations.
  4. If we generalize the Pearson family to the family of distributions characterized by the differential equation  $d(\ln f(u))/du = (a + u)/(b_0 + \sum_{i=1}^m b_i u^i)$  with higher order polynomials replacing the quadratic function, moments of higher orders up to  $m$  will be used.

### References

- Afriat, S. N. [1972], "Efficiency Estimation of Production Functions," International Economic Review 13, pp. 568-598.
- Aigner, D. J. and S. F. Chu [1968], "On Estimating the Industry Production Function," American Economic Review 58, pp. 826-835.
- Aigner, D. J., C. A. K. Lovell, and P. Schmidt [1977], "Formulation and Estimation of Stochastic Frontier Production Function Models," Journal of Econometrics 6, pp. 21-37.
- Aitchison, J. and S. D. Silvey [1958], "Maximum-Likelihood Estimation of Parameters Subject to Restraints," Annals of Mathematical Statistics 29, pp. 813-828.
- Bera, A. K. and C. M. Jarque [1980], "An Efficient Large-Sample Test for Normality and Homoscedasticity of Regression Residuals," Manuscript, Australian National University.
- Breusch, T. S. and A. R. Pagan [1980], "The Lagrange Multiplier Test and Its Applications to Model Specification in Econometrics," Review of Economic Studies XLVII, pp. 239-253.
- Broek, J. van den; F. R. Forsund; L. Hjalmarsson and W. Meeusen [1980], "On the Estimation of Deterministic and Stochastic Frontier Production Functions--A Comparison," Journal of Econometrics 13, pp. 117-138.
- Cohen, A. C., Jr. [1951], "Estimation of Parameters in Truncated Pearson Frequency Distributions," Annals of Mathematical Statistics 22, pp. 256-265.
- Elderton, W. P. and N. L. Johnson [1969], Systems of Frequency Curves, Cambridge University Press, Cambridge, London.
- Farrell, M. J. [1957], "The Measurement of Productive Efficiency," Journal of the Royal Statistical Society (A, general) 120, pp. 253-281.
- Forsund, F. R.; C. A. K. Lovell and P. Schmidt [1980], "A Survey of Frontier Production Functions and of Their Relationship to Efficiency Measurement," Journal of Econometrics 13, pp. 5-25.
- Greene, W. [1980], "Maximum Likelihood Estimation of Econometric Frontier Functions," Journal of Econometrics 13, pp. 27-56.
- Johnson, N. L. and S. Kotz [1970], Continuous Univariate Distributions--I, Houghton Mifflin Co., Boston.
- Lee, L. F. and W. G. Tyler [1978], "The Stochastic Frontier Production Function and Average Efficiency: An Empirical Analysis," Journal of Econometrics 7, pp. 385-389.

- Lee, L. F. [1981], "Consistent Estimation of a Multivariate Doubly Truncated or Censored Tobit Model," Manuscript, University of Florida.
- Lee, L. F. [1981], "Estimation of Some Non-Normal Limited Dependent Variable Models," Manuscript, University of Florida.
- Lee, L. F. [1981], "A Specification Test for Normality in the Generalized Censored Regression Models," Manuscript, University of Minnesota.
- Meeusen, W. and J. van den Broeck [1977], "Efficiency Estimation From Cobb-Douglas Production Functions with Composed Error," International Economic Review 18, pp. 435-444.
- Neyman, J. [1959], "Optimal Asymptotic Tests of Composite Statistical Hypotheses," in Probability and Statistics, ed. U. Grenander, pp. 213-234, New York: Wiley.
- Pitt, M. M. and L. F. Lee [1981], "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," forthcoming in Journal of Development Economics.
- Rao, C. R. [1973], Linear Statistical Inference and Its Applications, 2nd edition, John Wiley & Sons, New York.
- Richmond, J. [1974], "Estimating the Efficiency of Production," International Economic Review 15, pp. 515-521.
- Silvey, S. D. [1959], "The Lagrangean Multiplier Test," Annals of Mathematical Statistics 30, pp. 389-407.
- Stevenson, R. E. [1980], "Likelihood Functions for Generalized Stochastic Frontier Estimation," Journal of Econometrics 13, pp. 57-66.
- Timmer, C. P. [1971], "Using a Probabilistic Frontier Production Function to Measure Technical Efficiency," Journal of Political Economy 79, pp. 776-794.
- Tyler, W. G. and L. F. Lee [1979], "On Estimating Stochastic Frontier Production Functions and Average Efficiency: An Empirical Analysis with Colombian Micro Data," Review of Economics and Statistics 61, pp. 435-438.