POINT AND INTERVAL ESTIMATION IN CROSS-SECTIONAL STOCHASTIC FRONTIER MODELS: THE EFFECTS OF SAMPLE SIZE

by

Charles Adam Flynt

A Thesis Submitted to the Faculty of the
DEPARTMENT OF AGRICULTURAL AND RESOURCE ECONOMICS

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

2005

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of the requirements for an advanced degree at The University of Arizona.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment, the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained form the author.

SIGNED: ____________________

APPROVAL BY THESIS DIRECTORS

This thesis has been approved on the dates shown below:

Bruce R. Beattie
Professor
Agricultural and Resource Economics

5/13/05
Date

Satheesh Aradhyula
Professor
Agricultural and Resource Economics

5/13/05
Date

Acknowledgments

I wish to thank, first and foremost, Dr. Bruce Beattie, Dr. Gary D. Thompson and Dr. Satheesh Aradhyula for their constant encouragement and helpful suggestions for improving the quality of this thesis. Without their careful guidance, this work could never have been completed and I am greatly indebted to each of them.

I would also like to thank Jørgen Mortensen for his help in acquiring empirical data which, though not directly used in this thesis, certainly motivated the Monte Carlo experiments on the estimators themselves. I am also grateful to Jørgen for our many conversations in passing and for his support during all stages of the thesis.

I am greatly indebted to Nancy Smith for her avid support of both this thesis and my academic pursuits in general. Since the day I applied to the graduate program, Nancy has always encouraged me and been willing to go the extra mile on my behalf. She is an invaluable asset to the graduate program and a true friend.

I am grateful to the Department of Agricultural and Resource Economics for providing both my funding and an enjoyable, friendly environment in which to pursue my academic goals. I have truly enjoyed this two year journey and will always cherish the memory of my time spent in the graduate program.

Last but by no means least, I wish to thank my amazing wife Anette. She has always believed in me and never ceased in her support of my efforts. I could never have come this far without her by my side.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Abstract

This study utilizes Monte Carlo experiments on simulated data to study the effects of sample size on the empirical accuracy of both point and interval estimates of technical efficiency in cross-sectional stochastic frontier models. Also considered is the robustness of Coelli's (1995) test statistic for the presence of skewness. It is found that sensitivity to sample size varies by model as well as the relative amount of inefficiency present in the data. Furthermore, large amounts of inefficiency are not optimal in terms of interval estimation accuracy. Finally, results indicate that Coelli's asymptotic test statistic is robust in moderately small samples, though performance varies with the underlying distribution of inefficiency.

Chapter 1

Introduction

Stochastic frontier models have been prevalent in the literature since their simultaneous introductions by Aigner et al. (1977) and Meeusen and van den Broeck (1977). Though methods for constructing firm–level estimates of technical efficiency were not immediately introduced, the composed error structure was generally considered a more theoretically sound basis for modeling technical efficiency than previously employed deterministic frontier methods.

Shortly after the introduction of the original composed error specification of the stochastic frontier model, Jondrow et al. (1982) purposed a method to obtain point estimates of firm-level technical efficiency, relying on either the conditional mean or mode of the distribution of the inefficiency. Battese and Coelli (1988) later expanded upon this estimator using a higher order approximation of the conditional mean that would theoretically yield more accurate point estimates. Their estimator is now considered almost standard. Later work by Bera and Sharma (1999) developed an interval estimator based on the conditional distribution of the inefficiency, while Horrace and Schmidt (1996) introduced an interval estimator based on the distribution for which the Battese and Coelli point estimator is the mean.

Additionally, Coelli (1995) has developed a specification test for the stochastic frontier model, consisting of a test for skewness in the residuals of a simple OLS regression. As noted by Bera and Mallick (2002), a test for skewness should not be misconstrued as a test for the stochastic frontier model itself. Skewness is a necessary condition for the stochastic frontier model, but it is not a sufficient one. While Coelli's test statistic is incapable of directly assessing the validity of the stochastic frontier model, it does give the researcher insight into whether or not a critical assumption of the model appears true. That is to say that while we cannot validate the claim that the data constitute a recipe for the stochastic frontier model, we can at least determine if one of its key ingredients is present.

While useful, Coelli's test statistic is an asymptotic test, with properties defined only for large samples. Similar arguments can be made for the point and interval estimators, both of which rely strongly upon the distributional assumptions underlying the stochastic frontier model itself. Small sample properties are not straightforwardly derived and are thus not well understood in these models. This lack of knowledge brings into question the reliability of both point and interval estimators, as well as Coelli's asymptotically normal test statistic, for smaller samples. Furthermore, differing specifications of the stochastic frontier model, particularly with regards to the distribution of the inefficiency component, may lead to differential performance as sample size grows shrinks. It would therefore be of interest to examine the small sample properties of these estimators, as well as Coelli's test statistic, in at least the most frequently employed stochastic frontier specifications.

To this end, a Monte Carlo experiment is conducted wherein three of the most commonly used stochastic frontier models are estimated using simulated data. By varying the sample size and the amount of inefficiency simulated in the data, we attempt to draw at least qualitative conclusions about the effects such factors may have upon the performance of the point and interval estimators. Coelli's test statistic is also computed in all cases, in an effort to gauge its ability to detect skewness in the composed OLS residuals.

The remainder of the thesis proceeds as follows: chapter two gives a theoretical background and brief survey of the relevant literature on stochastic frontier models, while chapter three considers the details of empirical estimation. Chapter four outlines the design of the Monte Carlo experiments, data generation, numerical optimization and the measures with which the estimators are compared. Chapter five presents the simulation results and chapter six provides concluding remarks.

Chapter 2

Stochastic Frontier Models

The general rationale of modeling a production process in terms of a frontier, as opposed to a production function, is that it serves as an upper bound on productive capability of a profit maximizing producer. That is, a production frontier accounts for the fact that while producers seek to maximize their profits, they are not always successful. In the words of Kumbhakar and Lovell (2000, pg. 2) “Not all producers succeed in utilizing the minimum inputs required to produce the outputs they choose to produce, given the technology at their disposal. ... Consequently, not all producers succeed in maximizing the profit resulting from their production activities.” By measuring shortfalls from the production frontier, inference can be drawn about the efficiency with which producers allocate their inputs in the productive process.

To make the difference between production functions and production frontiers explicit, consider a typical production function of the general form

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

where, for the *i*-th firm, y_i denotes output, x_i represents a vector of inputs and β is a vector of parameters characterizing the production process. Here, the right hand side of the equation is presumed to be the *frontier* of possible production, assuming profit maximization on the part of the producer.

If we alternatively consider the addition of a random, two-sided (say normal) disturbance term, *v*, then the result is a stochastic production function model, given by

$$y_i = f(x_i, \beta) + v_i \tag{2.2}$$

If, on the other hand, we consider the addition of a positive, one-sided error term[1], *u*, we would obtain

$$y_i = f(x_i, \beta) - u_i \tag{2.3}$$

The result in the one-sided error case is a deterministic frontier model. Given the assumptions about *u*, production will lie on or below the *frontier* given by the right-hand side less the disturbance. That is, for values of *u* greater than or equal to zero,

$$y_i \leq f(x_i, \beta) \tag{2.4}$$

Notice that in the stochastic production function model there is no frontier to speak of, in the sense that deviations from the production function can be either positive or negative. This provides a sharp contrast to the deterministic frontier model, where only one-sided deviations are allowed. Figures 2.1 and 2.2 graphically illustrate the differences between these two specifications in a single input environment.

The notion of a *stochastic* frontier departs from the deterministic frontier formulation in the sense that it allows the entire frontier to become stochastic. Models of this type have existed in the literature for some time, stemming from early work in frontier production models by Farrel (1957), Aigner and Chu (1968), Seitz (1971) and Timmer (1971). Two later studies by Aigner et al. (1977) and Meeusen and van den Broeck (1977) simultaneously proposed the estimation of such models using a composed error term. Both formulations began with a normally distributed error to account for pure randomness and added a second, one-sided error to account for technical inefficiency. In particular, Aigner et al. (1977) suggested a (positive) half normal distribution for the

[1]Equivalently, *u* can take the form of a symmetric error term truncated (from below) at zero.

inefficiency. In this context, the model would take on the form,

$$y_i = f(x_i, \beta) + v_i - u_i \tag{2.5}$$

Here, *u* is the one-sided, half-normal error component and *v* is a normally distributed error. The motivation for using a half normal distribution to characterize the inefficiency is best summarized by Kumbhakar and Lovell (2002, p. 74), who base this distributional assumption upon "... the plausible proposition that the modal value of technical inefficiency is zero, with increasing values of technical inefficiency becoming increasingly less likely." They also note that this particular distributional assumption allows for straightforward derivation of the density for the composed errors, which is required to undertake empirical estimation.

Though theoretically justifiable and mathematically one of the most tractable, the half normal distribution for u_i is but one of several composed error specifications considered in the stochastic frontier literature. Aigner et al. (1977) also considered the exponential distribution as an alternative to the half normal in their introduction of the frontier model, while the simultaneous introduction by Meeusen and van den Broeck (1977) focused exclusively on the exponential distribution. Stevenson (1980) introduced a more general truncated normal distribution, allowing for a non-zero mode and thereby nesting the normal–half normal model of Aigner et al. Greene (1980) considered an even more flexible gamma distribution for the inefficiency component, u_i.

While it is tempting to employ a more flexible specification, Kumbhakar and Lovell (2000) point out several potential problems, stemming from research by Ritter and Simar (1997a, 1997b). These include difficulty in the estimation of the additional parameters, particularly in the case of the gamma distribution. Moreover, they argue that the log likelihood function for the gamma formulation is difficult (and in some cases impossible) to maximize. Kumbhakar and Lovell (2000) investigated results reported by Greene (1990) and found that rank correlations between pairs of firm-level technical efficiency estimates across all four models were highly correlated. They argue that such evidence supports the use of simpler distributions, namely the half normal and exponential, rather than the more general truncated normal and gamma counterparts. Furthermore, Kumbhakar and Lovell's results indicate that, at least insofar as the ranking of firm-level efficiencies, “...the choice between the two one-parameter [half normal and exponential] densities is largely immaterial” (p. 90).

Regardless of the choice in distributional specification for the inefficiency term, it should be clear that the production frontier is no longer deterministic but in fact stochastic, and is given for the *i*-th firm by

$$y_i \leq f(x_i, \beta) + v_i \qquad (2.6)$$

An important and perhaps subtle difference between the stochastic and deterministic specifications should be addressed at this point. Notice that in the aforementioned deterministic frontier model, no firm-level variation in output capability is allowed for after taking efficiency into account. The stochastic frontier model, by contrast, explicitly allows for such variation with the inclusion of *v*. That is, the specification of a stochastic frontier allows for random differences in operating environment that, while beyond the producer's control, are nonetheless related to productive capability. Indeed, Bera and Sharma (1999) exploit this very characteristic of the stochastic frontier framework to construct measures of production uncertainty. The deterministic frontier model simply assumes away any such influences on production.

Such an assumption is convenient and perhaps even appropriate in some settings, but it would seem in general to be an oversimplification.

In addition to their differing assumptions about the presence of statistical noise, there also exists a more fundamental distinction between the stochastic frontier framework and the deterministic approach. As highlighted by Horrace and Schmidt (1996), stochastic frontiers are statistical models and provide *estimates* of technical efficiency. Deterministic frontier models, by contrast, are often formulated as mathematical programming problems.[2] While uncertainty is accounted for and quantifiable in the stochastic frontier model, the deterministic frontier model produces only efficiency *measures*, about which uncertainty is not easily gauged. Indeed, Horrace and Schmidt (1996, p. 257) note that while confidence intervals for deterministic models can be generated through bootstrapping techniques, such procedures "...are an imperfect substitute for an adequately developed distributional theory."

While theoretically more appealing, the stochastic frontier model does introduce certain empirical obstacles not found in more simplified formulations. The root cause of these difficulties lies in the composed errors. In the present case of cross-sectional data, for example, the measurement of firm-level efficiency would seem to require the empirically impossible task of decomposing the observed errors into their inefficiency (u) and noise (v) components. In response to this apparent impasse, alternative techniques have been developed for extracting information on firm-level technical inefficiency. The following chapter will review these techniques in detail.

The stochastic frontier model assumes inefficiency to exist in a production

---

[2]There are techniques to estimate deterministic frontier models statistically, but these methods suffer from serious theoretical and empirical shortcomings. See Kumbhakar and Lovell (2000).

process characterized by some amount of stochastic variation beyond the control of the producer. In general, this appears to be an unreasonable assumption. Much less tenable would be the assumption that the productive process was purely deterministic, and that observed deviations were entirely attributable to inefficiency on the part of the producer. It is for this reason that the stochastic frontier framework is often adopted in favor of its more rigid deterministic counterpart.

Stochastic Production Function

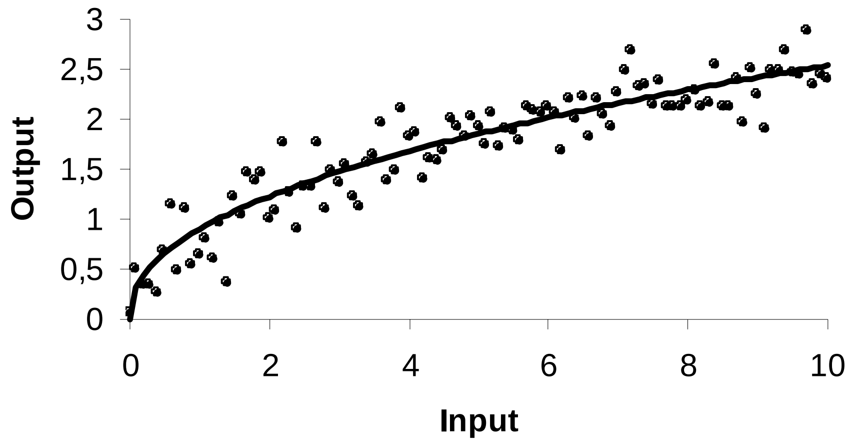


Figure 2.1

Deterministic Production Frontier

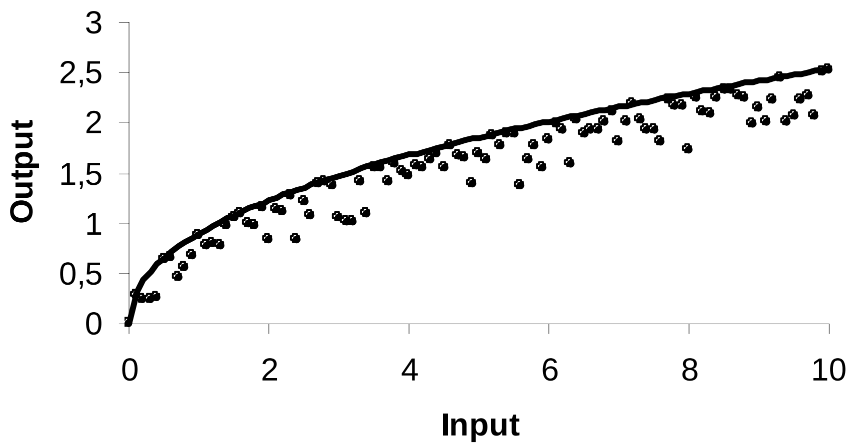


Figure 2.2

Chapter 3

Empirical Estimation of Technical Efficiency

Stochastic frontier models are composed error models, with one component of the error representing inefficiency and the other statistical noise. While simple enough in theory, such models present challenges in both estimation and inference. This chapter considers cross-sectional stochastic frontier models[3] under three of the most common error specifications. Section one explores techniques for estimation, extraction of firm-level technical efficiency estimates and the construction of confidence intervals for the normal – half normal model. Subsequent sections generalize these techniques to the normal – exponential and normal – truncated normal models.

3.1 Estimation: The Normal – Half Normal Model

Recall, from chapter 2, the basic stochastic production frontier model can be expressed as

$$y_i = x_i\beta + v_i - u_i \qquad i = 1, ..., n \qquad (3.1)$$

where, for the i-th firm, y_i typically denotes the logarithm of output, x_i denotes a k-vector of input quantities, also typically measured in logarithms (or functions of input quantities), β represents technology parameters to be estimated and v_i and u_i represent noise and inefficiency, respectively. Given some distributional assumptions on the components of the error terms, such models can be estimated by either maximum likelihood or (in some cases) a two-stage, corrected ordinary least squares (COLS) procedure. Consider, for example, the following set of assumptions:

[3] Panel data models also exist, but are beyond the scope of this analysis. For an excellent extension of the techniques in this chapter to panel data, see Kumbhakar and Lovell (2000).

(i) $v_i \underset{iid}{\sim} N(0,\ \sigma_v^2)$

(ii) $u_i \underset{iid}{\sim} \left|N(0,\ \sigma_u^2)\right|$

(iii) v_i and u_i are distributed independent of one another, as well as the regressors in x. Here, v is the standard, two-sided error and u represents the one-sided (non-negative) inefficiency. Under these assumptions, one obtains the so-called "Normal – Half-Normal" model of Aigner et al (1977).

To effect maximum likelihood or COLS estimation of this model, the distribution of the composed errors must be specified. Aigner et al. (1977) as well as Kumbhakar and Lovell (2000) derive this distribution by first appealing to assumption (iii), which suggests that given[4]

$$f(u) = \frac{1}{\sqrt{2\pi}\sigma_u \Phi(0/\sigma_u)} \exp\left\{-\frac{u^2}{2\sigma_u^2}\right\}$$
$$= \frac{2}{\sqrt{2\pi}\sigma_u} \exp\left\{-\frac{u^2}{2\sigma_u^2}\right\} \qquad \text{and}$$
$$f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{-\frac{v^2}{2\sigma_v^2}\right\}$$

the joint density function of v and u will simply be the product of the individual densities. That is

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

Denoting the composed error (v –

[4]Notice that $f(u)$ is distributed as a $N(0,\ \sigma_u)$ that is truncated from below at zero.

u) as ε, the joint distribution of *u* and ε can be obtained by direct substitution as

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

Again following Aigner et al. (1977) and Kumbhakar and Lovell (2000), the marginal density of ε is found by integrating out the density of *u*, so that

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

where $\lambda=\sigma_u/\sigma_v$ and $\sigma=(\sigma_u^2+\sigma_v^2)^{1/2}$

This density is plotted in figures 3.1 – 3.3 for several values of λ. The mean and variance of ε can be derived by considering the corresponding mean and variance of its *v* and *u* components. Kumbhakar and Lovell (2000) show that

$$\begin{aligned} E(\varepsilon) &= E(v)-E(u) \\ &= 0-E(u) \\ &= -\sigma_u\sqrt{\frac{2}{\pi}} \end{aligned}$$

$$\begin{aligned} V(\varepsilon) &= V(v-u) \\ &= V(v)+V(u)\text{ - }2Cov(v,\ u) \\ &= V(v)+V(u)+0 \\ &= \sigma_v^2+\frac{\pi\text{ - }2}{\pi}\sigma_u^2 \end{aligned}$$

Having obtained the density of the composed errors in this fashion, maximum likelihood estimation becomes straightforward. The likelihood function for a sample of *n* firms can be written as

$$L(\varepsilon_i,\lambda,\sigma)=\prod_{i=1}^{n}\frac{2}{\sigma}\phi\left\{\frac{\varepsilon_i}{\sigma}\right\}\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right)$$
$$=\left(\frac{2}{\sigma}\right)^n\left(\frac{1}{\sqrt{2\pi}}\right)^n\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^{\,2}\right\}\prod_{i=1}^{n}\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right)$$

The log likelihood function is then found by taking the natural logarithm, yielding

$$\ln L(\varepsilon_i,\lambda,\sigma)=\ln\left(\frac{2}{\sigma}\right)^n+\ln\left(\frac{1}{\sqrt{2\pi}}\right)^n-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^{\,2}+\sum_{i=1}^{n}\ln\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right)$$
$$=n\ln(2)-n\ln(\sigma)+n\ln\left(\frac{1}{\sqrt{2\pi}}\right)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^{\,2}+\sum_{i=1}^{n}\ln\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right)$$
$$=constant-n\ln(\sigma)-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\varepsilon_i^{\,2}+\sum_{i=1}^{n}\ln\Phi\left(-\frac{\varepsilon_i\lambda}{\sigma}\right)$$

In this formulation, λ and σ (or, equivalently, σ_u^2 and σ_v^2) become additional parameters in the model and can be consistently estimated along with the technology parameters in β[5]. Hence, there are *k*+2 parameters to estimate.

As an alternative to maximum likelihood, COLS can also be used to obtain consistent estimates of all parameters in the model. First, we re-write the stochastic production frontier as

$$y_i=\beta_0+\sum_{i=1}^{k}\beta_k x_{ki}+v_i-u_i$$

As was shown in (3.7), the composed residuals under inefficiency have a non-zero mean and so it is necessary to transform the model as follows:

$$y_i=\beta_0+E(v_i-u_i)+\sum_{i=1}^{k}\beta_k x_{ki}+v_i-u_i-E(v_i-u_i)$$
$$=\beta_0-E(u_i)+\sum_{i=1}^{k}\beta_k x_{ki}+v_i-u_i+E(u_i)$$
$$=\left[\beta_0-E(u_i)\right]+\sum_{i=1}^{k}\beta_k x_{ki}+\left\{v_i-u_i+E(u_i)\right\}$$

(3.11)

Reformulated in this manner, the error term has a zero mean and finite variance, allowing for consistent estimates of all parameters, apart from the intercept, by OLS. As can be readily seen, the resulting OLS intercept will be biased downward by the mean of *u*. In order to obtain a consistent estimate of this term, a consistent estimator for the mean of *u* is needed.

Given assumptions (i), (ii) and (iii), a consistent estimator for β_0 can be obtained by considering the central sample moments of the OLS residuals and the associated central moments of the error term. Denoting the *i*-th sample moment of the residuals about their mean m as m_i and the *i*-th moment of error term about its mean μ as μ_i, we have[6]

$$m = \frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i) \qquad \mu = \sigma_u\sqrt{(2/\pi)}$$

$$m_2 = \frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i - m)^2 \qquad \mu_2 = \sigma_v^2 + \frac{\pi-2}{\pi}\sigma_u^2$$

$$m_3 = \frac{1}{n}\sum_{i=1}^{n}(\varepsilon_i - m)^3 \qquad \mu_3 = \sqrt{\frac{2}{\pi}}\left(1-\frac{4}{\pi}\right)\sigma_u^3$$

Because the central sample moments of the residuals are asymptotically equivalent to the central moments of the error term, m_3 can be set equal to μ_3, implying

$$\hat{\sigma}_u^2 = \left(\sqrt{\frac{\pi}{2}}\frac{\pi}{\pi-4}m_3\right)^{2/3}$$

[6] The moments of the (OLS) error term are given by Olson et al. (1980) and Kumbhakar and Lovell (2000).

$$m_3 > 0 \qquad (3.13)$$

as a consistent estimator for the σ_u^2 parameter[7]. By considering this result and equating m_2 with μ_2 it is then trivial to obtain

$$\hat{\sigma}_v^2 = m_2 - \frac{\pi - 2}{\pi}\hat{\sigma}_u^2$$

as a consistent estimator for σ_v^2, the variance of v. Having found consistent estimators for all other parameters, a consistent COLS estimate of intercept term is then given by

$$\hat{\beta}_0 = \left[\beta_0 - E\left(u_i\right)\right] + E\left(u_i\right)$$
$$\hat{\beta}_0 = \left[\beta_0 - E\left(u_i\right)\right] + \hat{\sigma}_u\sqrt{(2/\pi)}$$
$$\hat{\beta}_0 = \text{OLS intercept} + \hat{\sigma}_u\sqrt{(2/\pi)}$$

COLS has the advantage of being robust to distributional assumptions about the disturbances during the first stage estimation of the technology parameters (excluding the constant), but it is not without its own shortcomings. The reliance on the moments of the residuals in estimating the variances of the disturbance terms, as well as for correcting the bias in the intercept, hinges critically upon the sign of the third central sample moment. Olson et al. (1980) observed that negativity of the third central moment of the error term does not ensure that the third central moment of the residuals will always be negative. In practice it is entirely possible to observe a non-negative, third central moment of the residuals, although this would bring into serious question whether or not the stochastic frontier model was correctly specified. In addition, COLS estimation of the stochastic frontier model has been found to be inefficient compared to maximum likelihood, a direct

[7] Notice that σ_u^2 is *not* the variance of u. Recall that $u \sim |N(0,\sigma_u^2)|$ so that $V(u) = ((\pi-2)/\pi)\,\sigma_u^2$. See Aigner et al. (1977).

consequence of the latter's usage of distributional assumptions (or, equivalently, the former's lack thereof). It is worth mentioning, however, that this difference in efficiency holds only for the case where the underlying distributional assumptions of the maximum likelihood specification are in fact correct.

When maximum likelihood estimation is used, Kumbhakar and Lovell (2000) point out that the re-parameterization to λ is particularly convenient, because it yields information about the composition of ε. Specifically, they note that as $\lambda \to \infty$ the one-sided error 'dominates' and a deterministic frontier model obtains. Conversely, as $\lambda \to 0$, a stochastic production function model with no inefficiency emerges. They also acknowledge that while the value of λ is useful in trying gauge the presence of inefficiency, it is difficult to interpret the results of a standard likelihood ratio or Wald test for the boundary values of the parameter space (e.g. $\lambda=0$). Because of this, inference about the presence of inefficiency is not easily made solely on the basis of λ.

Other methods of testing for the presence of inefficiency are available, however. One such test, suggested by Coelli (1995), does not require maximum likelihood estimation of the full model but instead only simple OLS regression of a restricted model. Because the composed residuals must exhibit negative skewness in the presence of inefficiency, Coelli proposed a relatively straight-forward method for testing the null hypothesis of zero skewness. Again denoting the *i*-th central sample moment of the residuals about their mean as m_i , his test statistic can be written as

$$Z = \frac{m_3}{(6m_2^3 / n)^{1/2}}$$

and is distributed asymptotically as a standard normal. For any intrinsically linear (e.g. Cobb Douglas, translog, etc.) model, this statistic can be easily computed using the residuals generated by a standard OLS regression. Because this test statistic follows an *asymptotically* normal distribution, however, inference in small sample applications may warrant some measure of caution.

While such testing can readily provide inference about the presence of technical inefficiency, it yields little *quantitative* insight. In order to say something about the extent of efficiency with which firms are operating, it is necessary to further examine the residuals of the full model. Following estimation by either maximum likelihood or COLS, however, one obtains only *composed* residuals. These residuals naturally contain information about efficiency, but it is obscured by statistical noise from which it cannot be readily separated. One early solution to this problem was the usage of an "average efficiency" measure. The first such estimator was proposed by Aigner et al (1977), and is given by

$$\begin{aligned} \mathrm{TE}_{\mathrm{ALS}} &= \left[1 - E(u)\right] \\ &= 1 - \sigma_u \sqrt{(2/\pi)} \end{aligned}$$

An alternative estimator,

$$\begin{aligned} \mathrm{TE}_{\mathrm{LT}} &= E\left(\exp\{-u\}\right) \\ &= 2\left[1 - \Phi\left(\sigma_u\right)\right] \exp\left\{\frac{\sigma_u^2}{2}\right\} \end{aligned}$$

was later proposed by Lee and Tyler (1978). Kumbhakar and Lovell (2000) note that the latter estimator is preferred, as the former contains only the first term in the power series expansion of $\exp\{-u\}$. Regardless of which estimator is ultimately chosen, one can easily obtain an average efficiency measure and draw limited inference about the technical efficiency of a given sample of firms. The problem with this, however, is that averages can mask large (or small) inefficiencies at the firm level. One might even say that as the variation in firm level efficiency of a given sample grows, average efficiency measures will provide an increasingly abstract characterization of the individual firms. For such reasons, it is desirable to obtain estimates not just of average efficiency across firms in a given sample, but measures of firm-level efficiency itself.

While it is not possible to directly decompose an observed residual ε_i into its v_i and u_i components, there are means for extracting the information that ε_i contains about u_i. Jondrow et al. (1982) proposed two methods, both relying on the conditional distribution of u given the composed residual ε. Specifically, they show the conditional distribution under assumptions (i), (ii) and (iii) to be

$$
\begin{aligned}
f(u \mid \varepsilon) &= \frac{f(u,\varepsilon)}{f(\varepsilon)} \\
&= \frac{1}{\sqrt{2\pi}\sigma_*} \exp\left\{ -\frac{(u-\mu_*)^2}{2\sigma_*^2} \right\} \Big/ \left[1 - \Phi\left(-\frac{\mu_*}{\sigma_*} \right) \right]
\end{aligned}
$$

where $\mu_* = -\varepsilon\sigma_u^2 / \sigma^2$ and $\sigma_*^2 = \sigma_u^2\sigma_v^2 / \sigma^2$

Noting that one over the trailing denominator is the probability of a $N(\mu_*, \sigma_*)$ taking a positive value, the entire expression can be interpreted as the density of a $N(\mu_*, \sigma_*)$

truncated (from below) at zero.

Having the conditional density of *u* given ε, it is straightforward to obtain the conditional density of u_i for each firm in the sample given an estimate of its associated residual ε_i. Jondrow et al. (1982) proposed either the mean or mode of each firm's conditional density as a point estimate for the unobserved inefficiency error component u_i. Maintaining assumptions (i), (ii) and (iii) these estimators are, respectively,

$$E(u_i \mid \varepsilon_i) = \mu_{*i} + \sigma_* \left[\frac{\phi(-\mu_{*i}/\sigma_*)}{1-\Phi(-\mu_{*i}/\sigma_*)} \right]$$

and

$$M(u_i \mid \varepsilon_i) = \begin{cases} -\varepsilon\sigma_u^2/\sigma^2 & \text{if } \varepsilon_i \le 0 \\ 0 & \text{otherwise.} \end{cases}$$

Using either of these methods to obtain a point estimate of u_i, the Jondrow et al. (1982) estimator firm-level technical efficiency is then given by

$$TE_i = \exp\{-\hat{u}_i\} \qquad (3.$$

Battese and Coelli (1988) later introduced a second estimator,

$$\begin{aligned} TE_i &= E\left(\exp\{-u_i\} \mid \varepsilon_i\right) \\ &= \left[\frac{1-\Phi(\sigma_* - \mu_{*i}/\sigma_*)}{1-\Phi(-\mu_{*i}/\sigma_*)} \right] \exp\left\{-\mu_{*i} + \frac{1}{2}\sigma_*^2\right\} \end{aligned}$$

which is generally preferred for the same reason that Lee and Tyler's (1978) average efficiency estimator is preferred over that of Aigner et al (1977).

Using either of these estimators in conjunction with the conditional mean or mode of *u*, firm-level estimates of efficiency can be readily obtained. As Kumbhakar and

Lovell (2000) point out, however, these estimates are inconsistent8 in the cross-sectional context regardless of how one choses to estimate u_i and TE_i. As an alternative to relying solely on point estimates, Horrace and Schmidt (1996) have derived interval estimators for the normal – half normal model. Specifically, they exploit the fact that $u_i|\varepsilon_i$ follows a truncated normal distribution and thereby derive its upper and lower confidence bounds. Since $\exp\{-u_i\}$ is a monotonic transformation of u_i, they are then able to derive upper and lower bounds on the distribution of $\exp\{-u_i\}|\varepsilon_i$, for which the Battese and Coelli estimator is the mean. At a significance level of α, the upper and lower bounds on $\exp\{-u_i\}|\varepsilon_i$ are given by

$$LB = \exp\left(-\mu_{*i} - z_L \sigma_*\right)$$
$$UB = \exp\left(-\mu_{*i} - z_U \sigma_*\right)$$

where

$$z_L = \Phi^{-1}\left\{1-(\alpha/2)\left[1-\Phi\left(-\mu_{*i}/\sigma_*\right)\right]\right\}$$
$$z_U = \Phi^{-1}\left\{1-(1-\alpha/2)\left[1-\Phi\left(-\mu_{*i}/\sigma_*\right)\right]\right\}$$

By using these confidence intervals in conjunction with the Battese and Coelli (1988) point estimator, one can make inferences about firm-level technical efficiency while also

[8]This is a result of the variance of $(u_i|\varepsilon_i)$ being independent of i and thus not necessarily shrinking as sample size grows. See Kumbhakar and Lovell (2000).

knowing something about the uncertainty which surrounds those point estimates. Horrace and Schmidt (1996) note that the conditional distribution of u_i is obtained with the implicit assumption that the parameters β, λ and σ are known. Consequently, the confidence intervals for the distribution of $u_i|\varepsilon_i$, and therefore those of $\exp\{-u_i\}|\varepsilon_i$, characterize only the uncertainty associated with point estimates made from within these distributions. Horrace and Schmidt (1996, p. 262) also argue, however, that "For large *N* this is probably unimportant, since the variability in the parameter estimates is small compared to the variability intrinsic to the distribution of $u_i|\varepsilon_i$."

3.2 Estimation: The Normal – Exponential Model

Though it is not frequently employed in empirical work, the one parameter exponential distribution can also be used to characterize the inefficiency component of the stochastic frontier model given in section 3.1. In this case, the following assumptions are made:

Just as in the normal – half normal case, estimation can be carried out using either maximum likelihood or COLS. The required distribution of the composed errors can be derived analogously to those of section 3.1. We have

$$f(u) = \frac{1}{\sigma_u}\exp\left\{-\frac{u}{\sigma_u}\right\} \qquad \text{and}$$

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma_v}\exp\left\{-\frac{v^2}{2\sigma_v^2}\right\}$$

Given assumption (iii), the joint distribution is

$$\begin{aligned} f(u,v) &= \frac{1}{\sigma_u}\frac{1}{\sqrt{2\pi}\sigma_v}\exp\left\{-\frac{u}{\sigma_u}\right\}\exp\left\{-\frac{v^2}{2\sigma_v^2}\right\} \\ &= \frac{1}{\sqrt{2\pi}\sigma_u\sigma_v}\exp\left\{-\frac{u}{\sigma_u}-\frac{v^2}{2\sigma_v^2}\right\} \end{aligned}$$

(3.28)

By substitution of (ε + *u*) for *v* one obtains,

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

As shown in Aigner et al. (1977) and Kumbhakar and Lovell (2000), the density of the composed errors is found by integrating *u* out of the joint distribution. That is,

$$\begin{aligned} f(\varepsilon) &= \int_0^\infty \frac{1}{\sqrt{2\pi}\sigma_u\sigma_v}\exp\left\{-\frac{u}{\sigma_u} - \frac{(\varepsilon+u)^2}{2\sigma_v^2}\right\} du \\ &= \left(\frac{1}{\sigma_u}\right)\Phi\left(-\frac{\varepsilon}{\sigma_v} - \frac{\sigma_v}{\sigma_u}\right)\exp\left\{\frac{\varepsilon}{\sigma_u} + \frac{\sigma_v^2}{2\sigma_u^2}\right\} \\ &= \left(\frac{1}{\sigma_u}\right)\Phi(-A)\exp\left\{\frac{\varepsilon}{\sigma_u} + \frac{\sigma_v^2}{2\sigma_u^2}\right\} \end{aligned}$$

where $A = -\tilde{\mu}/\sigma_v$ and $\tilde{\mu} = -\varepsilon - (\sigma_v^2/\sigma_u)$

Figures 3.4 – 3.6 show this distribution of ε for various values of λ. Given this distribution, maximum likelihood estimation once more becomes straightforward. As in section 3.1, the likelihood function for a sample of *n* firms can be written as

$$\begin{aligned} L(\varepsilon_i, \sigma_u^2, \sigma_v^2) &= \prod_{i=1}^{n}\left(\frac{1}{\sigma_u}\right)\Phi(-A_i)\exp\left\{\frac{\varepsilon_i}{\sigma_u} + \frac{\sigma_v^2}{2\sigma_u^2}\right\} \\ &= \left(\frac{1}{\sigma_u}\right)^n \exp\left\{\sum_{i=1}^{n}\frac{\varepsilon_i}{\sigma_u} + \sum_{i=1}^{n}\frac{\sigma_v^2}{2\sigma_u^2}\right\}\prod_{i=1}^{n}\Phi(-A_i) \end{aligned}$$

Taking the natural logarithm yields the log likelihood function

$$\begin{aligned} \ln L(\varepsilon_i, \sigma_u^2, \sigma_v^2) &= \ln\left(\frac{1}{\sigma_u}\right)^n + \sum_{i=1}^{n}\frac{\varepsilon_i}{\sigma_u} + \sum_{i=1}^{n}\frac{\sigma_v^2}{2\sigma_u^2} + \sum_{i=1}^{n}\ln\Phi(-A_i) \\ &= -n\ln(\sigma_u) + \sum_{i=1}^{n}\frac{\varepsilon_i}{\sigma_u} + n\left(\frac{\sigma_v^2}{2\sigma_u^2}\right) + \sum_{i=1}^{n}\ln\Phi(-A_i) \end{aligned}$$

(3.32)

Notice that in this particular formulation, there is no re-parameterization in terms of λ and σ , as was used in the normal – half normal model. There remain k+2 parameters to be estimated, however, just as in the normal – half normal case.

Estimation of technical efficiency in the exponential model follows directly from the techniques used in section 3.1. A measure of average technical efficiency can be obtained from

$$
\begin{aligned}
TE_{ALS} &= \left[1 - E(u)\right] \\
&= 1 - \sigma_u
\end{aligned}
$$

The Jondrow et al. (1982) and Battese and Coelli (1988) point estimators can be found in a similar fashion to those of the normal – half normal model. As shown[9] in Kumbhakar and Lovell (2000), the conditional distribution of $u_i|\varepsilon_i$ is given by

$$
\begin{aligned}
f(u \mid \varepsilon) &= \frac{f(u,\varepsilon)}{f(\varepsilon)} \\
&= \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{-\frac{(u-\tilde{\mu})^2}{2\sigma_v^2}\right\} \Bigg/ \left[1-\Phi\left(-\frac{\tilde{\mu}}{\sigma_v}\right)\right]
\end{aligned}
$$

Because $u_i|\varepsilon_i$ takes the form of a N(μ, σ_v) truncated (from below) at zero, a straightforward extension of the Jondrow et al. technique can be applied, using either the mean or mode as a point estimate of u_i. These estimators, given in Kumbhakar and Lovell (2000), are

[9]The errors found in Kumbhakar and Lovell (2000, pg. 82) have been corrected.

$$E(u_i \mid \varepsilon_i) = \tilde{\mu}_i + \sigma_v \left[\frac{\phi\left(-\tilde{\mu}_i/\sigma_v\right)}{1-\Phi\left(-\tilde{\mu}_i/\sigma_v\right)} \right]$$

and

$$M(u_i \mid \varepsilon_i) = \begin{cases} \tilde{\mu}_i & \text{if } \tilde{\mu}_i \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

As with the normal – half normal case, these point estimates of u_i can be translated into point estimates of technical efficiency using the original estimator of Aigner et al. (1977),

(3.

$$TE_i = \exp\left\{-\hat{u}_i\right\}$$

Alternatively, the Battese and Coelli (1988) estimator can be readily adapted to the normal – exponential model and is given by

$$\begin{aligned} TE_i &= E\left(\exp\left\{-u_i\right\} \mid \varepsilon_i\right) \\ &= \left[\frac{1-\Phi\left(\sigma_v - \tilde{\mu}_i/\sigma_v\right)}{1-\Phi\left(-\tilde{\mu}_i/\sigma_v\right)} \right] \exp\left\{-\tilde{\mu}_i + \frac{1}{2}\sigma_v^2\right\} \end{aligned}$$

Likewise, the interval estimators of Horrace and Schmidt (1996) can be modified so that, at a significance level of α, the upper and lower bounds of $\exp\{-u_i\}|\varepsilon_i$ become

$$\begin{aligned} LB &= \exp\left(-\tilde{\mu}_i - z_L \sigma_v\right) \\ UB &= \exp\left(-\tilde{\mu}_i - z_U \sigma_v\right) \end{aligned}$$

where

$$\begin{aligned} z_L &= \Phi^{-1}\left\{1-(\alpha/2)\left[1-\Phi\left(-\tilde{\mu}_i/\sigma_v\right)\right]\right\} \\ z_U &= \Phi^{-1}\left\{1-(1-\alpha/2)\left[1-\Phi\left(-\tilde{\mu}_i/\sigma_v\right)\right]\right\} \end{aligned}$$

3.3 Estimation: The Normal – Truncated Normal Model

The normal – truncated normal model is a more general form of the normal – half normal model, allowing for a non-zero mean, μ , in the distribution of inefficiency. Here, the distributional assumptions of the model are:

As shown in Kumbhakar and Lovell (2000) and Stevenson (1980), the distribution of the composed errors in this model can be obtained in an identical fashion to those of the half normal and exponential models. We have

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

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} \exp\left\{-\frac{v^2}{2\sigma_v^2}\right\}$$

which, by assumption (iii), implies

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

Making the substitution of (ε + *u*) for *v* once more, we have

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

Following either Kumbhakar and Lovell (2000) or Stevenson (1980), the density of the composed errors is found by integrating *u* out to obtain

$$f(\varepsilon) = \int_0^\infty \frac{1}{2\pi\sigma_u\sigma_v\Phi(\mu/\sigma_u)} \exp\left\{-\frac{(u-\mu)^2}{2\sigma_u^2} - \frac{(\varepsilon+u)^2}{2\sigma_v^2}\right\} du$$
$$= \frac{1}{\sqrt{2\pi}\sigma\Phi(\mu/\sigma_u)} \Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon\lambda}{\sigma}\right) \exp\left\{-\frac{(\varepsilon+u)^2}{2\sigma^2}\right\}$$
$$= \left(\frac{1}{\sigma}\right)\phi\left(\frac{\varepsilon+\mu}{\sigma}\right)\Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon\lambda}{\sigma}\right)\left[\Phi\left(\frac{\mu}{\sigma_u}\right)\right]^{-1}$$

(3.45)

This distribution is shown in figures 3.7 – 3.9 for several values of λ. At this point, maximum likelihood estimation can be carried out in an identical fashion to the half normal and exponential models. For a sample of *n* firms the likelihood function is

$$L(\varepsilon_i,\sigma_u^2,\sigma_v^2,\mu)=\prod_{i=1}^{n}\left(\frac{1}{\sigma}\right)\phi\left(\frac{\varepsilon_i+\mu}{\sigma}\right)\Phi\left(\frac{\mu}{\sigma\lambda}-\frac{\varepsilon_i\lambda}{\sigma}\right)\left[\Phi\left(\frac{\mu}{\sigma_u}\right)\right]^{-1}$$
$$=\left(\frac{1}{\sigma}\right)^n\left[\Phi\left(\frac{\mu}{\sigma_u}\right)\right]^{-n}\left(\frac{1}{\sqrt{2\pi}}\right)^n\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\varepsilon_i+\mu\right)^2\right\}$$
$$\times\prod_{i=1}^{n}\Phi\left(\frac{\mu}{\sigma\lambda}-\frac{\varepsilon_i\lambda}{\sigma}\right)$$

Taking the natural logarithm, the log likelihood function becomes

$$
\begin{aligned}
\ln L(\varepsilon_i, \sigma_u^2, \sigma_v^2, \mu) = \ln\left(\frac{1}{\sigma}\right)^n + \ln\left[\Phi\left(\frac{\mu}{\sigma_u}\right)\right]^{-n} + \ln\left(\frac{1}{\sqrt{2\pi}}\right)^n - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\varepsilon_i + \mu\right)^2 \\
+ \sum_{i=1}^{n} \ln\Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon_i\lambda}{\sigma}\right) \\
= -n\ln(\sigma) - n\ln\left[\Phi\left(\frac{\mu}{\sigma_u}\right)\right] + n\ln\left(\frac{1}{\sqrt{2\pi}}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\varepsilon_i + \mu\right)^2 \\
+ \sum_{i=1}^{n} \ln\Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon_i\lambda}{\sigma}\right) \\
= -n\ln(\sigma) - n\ln\left[\Phi\left(\frac{\mu}{\sigma_u}\right)\right] - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(\varepsilon_i + \mu\right)^2 + \sum_{i=1}^{n} \ln\Phi\left(\frac{\mu}{\sigma\lambda} - \frac{\varepsilon_i\lambda}{\sigma}\right)
\end{aligned}
\tag{3.47}
$$

Estimation of technical efficiency in the truncated normal model is accomplished in the same fashion as the half normal and exponential cases, with the exception that there are now k+3 parameters to be estimated. For brevity, we consider only the firm-level estimators of technical efficiency and not average measures. The interested reader is referred to Kumbhakar and Lovell (2000) for the analog to the average measures given in the previous sections.

To obtain firm level estimates of technical efficiency in the truncated normal model, Kumbhakar and Lovell (2000) show that

$$f(u \mid \varepsilon) = \frac{f(u,\varepsilon)}{f(\varepsilon)}$$
$$= \frac{1}{\sqrt{2\pi}\sigma_*} \exp\left\{ -\frac{(u-\tilde{\mu})^2}{2\sigma_*^2} \right\} \Bigg/ \left[1 - \Phi\left(-\frac{\tilde{\mu}}{\sigma_*^2} \right) \right]$$

where $\tilde{\mu} = \left(-\varepsilon\sigma_u^2 + \mu\sigma_v^2\right) \big/ \sigma^2$ and $\sigma_*^2 = \sigma_u^2\sigma_v^2 / \sigma^2$

which, as in the other two cases, can be seen as a N(μ, σ*) truncated below zero. As a result, the Jondrow et al. point estimators of u_i are simply given by

$$E(u_i \mid \varepsilon_i) = \tilde{\mu}_i + \sigma_* \left[\frac{\phi\left(-\tilde{\mu}_i / \sigma_*\right)}{1 - \Phi\left(-\tilde{\mu}_i / \sigma_*\right)} \right]$$

and

$$M(u_i \mid \varepsilon_i) = \begin{cases} \tilde{\mu}_i & \text{if } \tilde{\mu}_i \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Either of these estimates can be used to obtain the Jondrow et al. point estimate of technical efficiency using

$$TE_i = \exp\left\{-\hat{u}_i\right\} \qquad (3.$$

Just as with the half normal and exponential cases, the Battese and Coelli (1988) estimator can be adapted to the truncated normal model. Kumbhakar and Lovell (2000) show this to be

$$TE_i = E\left(\exp\{-u_i\} \mid \varepsilon_i\right)$$
$$= \left[\frac{1 - \Phi\left(\sigma_* - \tilde{\mu}_i / \sigma_*\right)}{1 - \Phi\left(-\tilde{\mu}_i / \sigma_*\right)} \right] \exp\left\{ -\tilde{\mu}_i + \frac{1}{2}\sigma_*^2 \right\}$$

Finally, the interval estimators of Horrace and Schmidt (1996) for the distribution of exp{-u_i}|ε_i can be adapted to the truncated normal model. At a significance level of α, the upper and lower bounds are given by

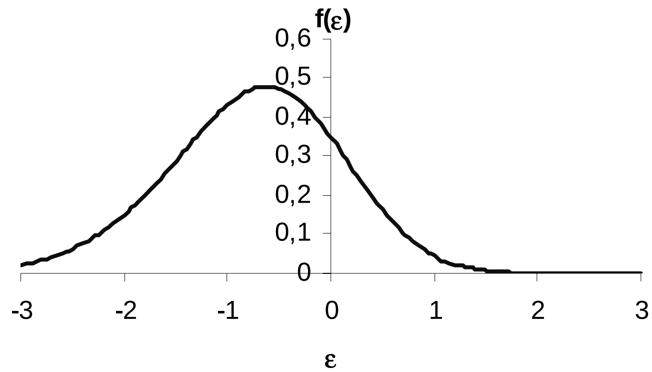
$$LB = \exp\left(-\tilde{\mu}_i - z_L \sigma_*\right)$$
$$UB = \exp\left(-\tilde{\mu}_i - z_U \sigma_*\right)$$

where

$$z_L = \Phi^{-1}\left\{1-(\alpha/2)\left[1-\Phi\left(-\tilde{\mu}_i/\sigma_*\right)\right]\right\}$$
$$z_U = \Phi^{-1}\left\{1-(1-\alpha/2)\left[1-\Phi\left(-\tilde{\mu}_i/\sigma_*\right)\right]\right\}$$

As can be clearly seen, the techniques applied to obtain both point and interval estimates, while somewhat complicated, are nonetheless general. Indeed, the same techniques can be applied to obtain point and interval estimates for other distributional specifications, including Greene's (1980) normal - gamma model. Furthermore, these same techniques can be adapted readily to panel data, though such exposition clearly extends beyond the scope of the present analysis.

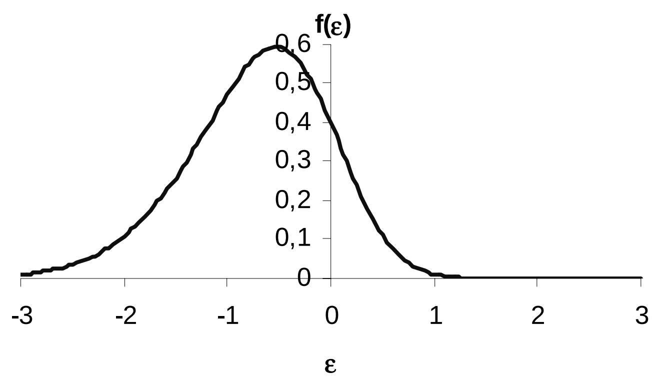
Normal – Half Normal



$\lambda = 1.5$

Figure 3.1

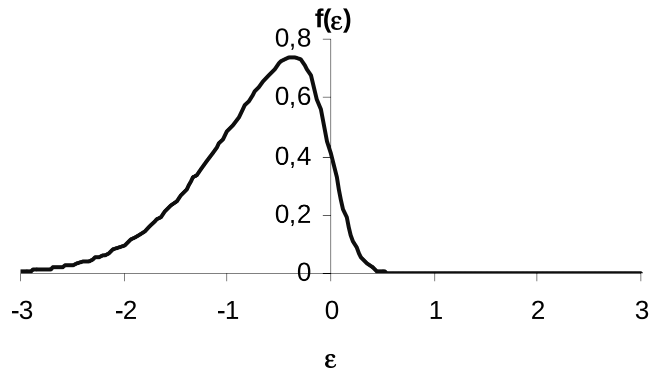
Normal – Half Normal



$\lambda = 2$

Figure 3.2

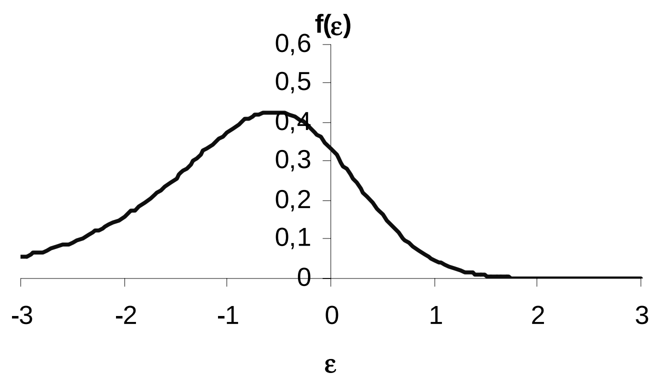
Normal – Half Normal



λ = 4.743

Figure 3.3

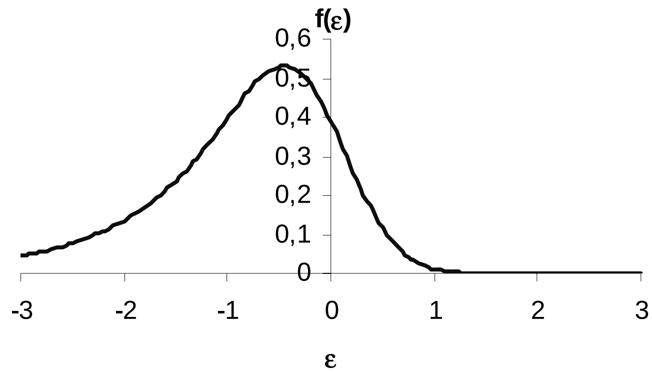
Normal – Exponential



λ = 1.5

Figure 3.4

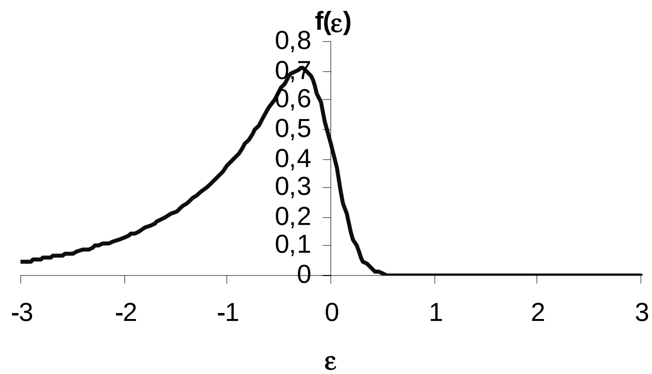
Normal – Exponential



λ = 2

Figure 3.5

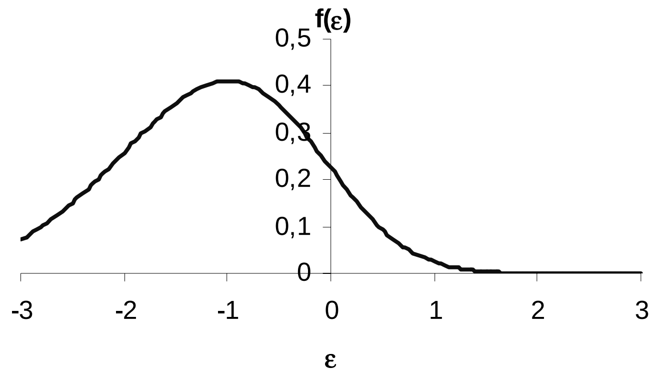
Normal – Exponential



λ = 4.743

Figure 3.6

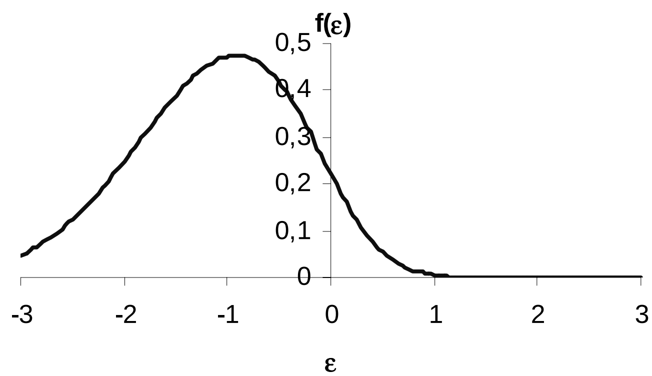
Normal – Truncated Normal



λ = 1.5

Figure 3.7

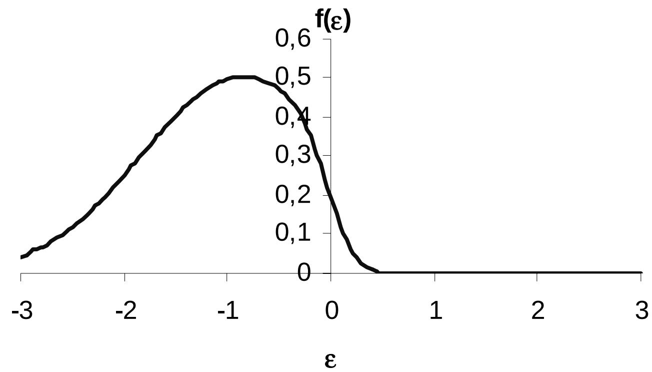
Normal – Truncated Normal



λ = 2

Figure 3.8

Normal – Truncated Normal



$\lambda = 4.743$

Figure 3.9

Chapter 4

Monte Carlo Experiments

This chapter provides details of the methods by which the Monte Carlo experiments in this study were undertaken. The first section outlines the type of simulations conducted; section two considers the methods used to generate the data for the simulations; section three briefly discusses estimation.

4.1 Simulation Design

The goal of this analysis is to better understand the effects of both sample size and the amount of inefficiency on the robustness of both point and interval estimators of firm-level technical efficiency. Also considered is the effect on Coelli's (1995) test for skewness.

The first experiment considers the case where each of the models is correctly specified and the data generating process corresponds exactly with the assumptions of the model. The normal – half normal, normal – exponential and normal – truncated normal model are each estimated by maximum likelihood, and the resulting parameter estimates are then used to construct the point and interval estimators of technical efficiency. Simulations are conducted with three levels of inefficiency, characterized by the size of the parameter λ, and across a series of sample sizes ranging from 50 to 500 observations. Coelli's (1995) test statistic is also computed using the residuals of a simple OLS regression.

The second experiment is concerned with the misspecification between the normal

– half normal and normal – exponential models. In this case, the misspecification considered is the use of the normal – half normal model when the data generating process is normal – exponential. The converse, where the data generating process is normal – half normal and the exponential model is applied, is also considered. Simulations are conducted across the same inefficiency levels and samples sizes as the first experiment, with both models estimated by maximum likelihood.

In both experiments, measures of parameter bias and mean square error (MSE) are calculated for σ_v^2, σ_u^2 and, where applicable, μ. These measures are constructed such that

$$BIAS = \left[\frac{1}{r}\sum_{i=1}^{r}\hat{\theta}_i\right] - \theta \qquad \text{and}$$

$$MSE = \frac{1}{r}\sum_{i=1}^{r}\left(\hat{\theta}_i - \theta\right)^2$$

where $\theta = \left\{\sigma_u^2,\ \sigma_v^2,\ \mu\right\}$ and r is the number of replications.

In the case of the point estimators of technical efficiency, performance is measured somewhat analogously to the parameter estimates by considering an average mean square error across replications. This can be expressed as

$$Avg.\ MSE = \frac{1}{r}\sum_{i=1}^{r}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\hat{u}_i - u_i\right)^2\right]$$

and is calculated for both the Jondrow et al. (1982) and Battese and Coelli (1988) point estimators.

In order to assess the empirical accuracy of the interval estimators of technical efficiency, we consider an average coverage rate, constructed as

$$Avg.\ Coverage = \frac{1}{r}\sum_{i=1}^{r}\left[\frac{1}{n}\sum_{i=1}^{n}\left(\delta_i\right)\right]$$

where $$\delta_i = \begin{cases} 1 & \text{if } LB_i < u_i < UB_i \\ 0 & \text{otherwise.} \end{cases}$$

Finally, with regards to Coelli's (1995) normally distributed test statistic for skewness of the OLS residuals, we calculate the average p-value across replications as a general metric for its performance.

4.2 Data Generation

For simplicity, all simulations conducted in this analysis employed a two-factor Cobb-Douglas production function, with the resulting stochastic frontier taking the form

$$y_i = Ax_{1i}^{\beta_1} x_{2i}^{\beta_2} e^{(v_i)} e^{(-u_i)}$$

In logarithms, this model can expressed equivalently as

$$\begin{aligned} \ln y_i &= \ln A + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i} + v_i - u_i \\ &= \beta_0 + \beta_1 \ln x_{1i} + \beta_2 \ln x_{2i} + v_i - u_i \end{aligned}$$

In this form the data can be easily generated using a variety of techniques. In the present analysis, rather than generating y_i itself, we use the second formulation to generate $\ln y_i$ using fixed[10] parameters β_0, β_1, and β_2, a matrix of stochastic $\ln(x_i)$ values and appropriately specified errors v_i and u_i. For every replication, $\ln x_{1i}$ and $\ln x_{2i}$ are drawn randomly from, respectively, a $N(10, 9)$ and $N(20, 36)$[11]. The v_i were generated for each

[10]Arbitrary values of, respectively, 10, 3 and 1 were chosen.

[11]These values were chosen arbitrarily, with the intent of generating random draws that were most often positive. Notice that negative values are still allowable, as these merely represent x values on the (0,1)

replication using random draws from a $N(0, \sigma_v^2)$, while the generation of u_i depend upon the form of inefficiency.

For the half-normal specification, the absolute values of random draws from a $N(0, \sigma_u^2)$ were used to simulate u_i. Similarly, for the truncated normal specification, absolute values were taken from random variates generated as $N(\mu, \sigma_u^2)$, where μ is non-zero mean. Finally, in the case of the exponential model, random draws from a uniform distribution were transformed[12] to random exponential draws according to the probability integral transformation. This transformation is given by

$$x = -\sigma_u \ln(1-U)$$

where U is distributed uniformly on the (0, 1) interval and the resulting x is distributed exponentially with mean σ_u and variance σ_u^2. The values assigned to σ_v^2 , σ_u^2 and μ during the experiments, as well as descriptive statistics on the composed errors, are detailed in figure 4.1.

4.3 Maximum Likelihood Estimation

For each simulation, maximum likelihood parameter estimates are found by iterative numerical optimization. Initial simulations employed a Quasi-Newton optimization routine but this was found in many cases to provide weaker[13] convergence than the non-

---

interval.

[12]Attempts were made to directly sample from random exponential draws, but SAS was found to be generating these variates with a mean and variance that differed significantly from σ_u and σ_u^2 ,respectively.

[13]Elements of the gradient under Quasi-Newton optimization frequently exceeded $1\text{x}10^{-3}$. A comparison of log likelihood values also showed the Simplex method to be obtaining superior optima with lower log likelihood values and smaller elements of the gradient.

derivative Nelder-Mead simplex algorithm. In order to obtain the best parameter estimates, the Quasi-Newton method was abandoned and the Nelder-Mead simplex method was employed exclusively.

In every case parameters were estimated directly, rather than re-parameterizing the models and estimating λ and σ. This was done primarily for the purposes of obtaining comparable estimates across models, because the exponential model has no direct analog to the σ parameter used in the other formulations. All calculations, ranging from parameter estimation to point and interval construction, were performed using SAS version 8.2.

Normal - Half Normal

	$\lambda = 1.5$		$\lambda = 2$		$\lambda = 4.743$	
Parameter	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Value	0.4	0.9	0.2	0.8	0.04	0.9
E($\mathcal{E}$)	-0.75694		-0.71365		-0.75694	
V($\mathcal{E}$)	0.72704		0.49070		0.36704	

Normal - Exponential

	$\lambda = 1.5$		$\lambda = 2$		$\lambda = 4.743$	
Parameter	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Value	0.4	0.9	0.2	0.8	0.04	0.9
E($\mathcal{E}$)	-0.94868		-0.89443		-0.94868	
V($\mathcal{E}$)	1.3		1		0.94	

Normal - Truncated Normal

	$\lambda = 1.5$			$\lambda = 2$			$\lambda = 4.743$		
Parameter	σ_v^2	σ_u^2	μ	σ_v^2	σ_u^2	μ	σ_v^2	σ_u^2	μ
Value	0.4	0.9	0.85	0.2	0.8	0.85	0.04	0.9	0.85
E($\mathcal{E}$)	-0.83246			-0.78664			-0.83246		
V($\mathcal{E}$)	0.90782			0.67018			0.54782		

Figure 4.1

Chapter 5

Simulation Results

This chapter summarizes the results of the two Monte Carlo experiments. Section one considers the case for which the true data generating process is known and appropriately modeled, while section two considers the case where the true data generating process differs from the assumptions of the model. Motivation for the simulations, again, lies in the fact that small sample properties for these point and interval estimators is largely unknown, as are the properties of Coelli's test statistic across various composed error specifications.

5.1 Simulation Results: Experiment One

The first experiment considers the case in which the models are correctly specified and proceeds to estimate them on a variety of sample sizes and levels of inefficiency, characterized by the magnitude of the parameter λ. The results of these simulations are presented in tables 1 – 5.

Some general patterns are readily noticeable in these results. Considering first the estimation of σ_v^2, σ_u^2 and, where applicable, μ, the effect of λ can be readily seen. In samples of every size and across all three models, the reduction in parameter bias and mean square error (MSE) as λ increases is pronounced. Nowhere is this more apparent than in the case of the normal – truncated normal specification, particularly with regards to the estimation of the placement parameter μ. While the effect of increasing λ is clear and consistent, the effect of sample size on bias and MSE is not. When the sample size is reduced from 500 to 350, for example, there is at least one instance in each model where parameter bias is seen to decline (in absolute value). In the case of the normal – truncated normal specification, there are two cases of a decrease in MSE when λ is at its smallest simulated value of 1.5. Though these are the only instances of MSE decreasing with sample size, the inconsistent changes in parameter bias can also be observed at least once in all three models when further reducing the sample size to 200. In the case of the normal – half normal and normal – exponential models, however, it appears that there is no discernible relationship between these somewhat counter intuitive results and the value of λ. While clearly inconsistent in general, the trend towards larger biases and MSE does hold consistently for the normal – half normal and normal – exponential models, regardless of the value of λ, for samples sizes of 200 and smaller.

In comparing bias and MSE across models, it is obvious that a large disparity exists between the normal – truncated normal and both the normal – half normal and normal – exponential models. This is true for all three values of λ and across all sample sizes, but is clearly the most evident when λ is small. Indeed, there is not a single case in which the empirical accuracy of parameter estimates from the normal – truncated normal specification exceeds those of the normal – half normal or normal – exponential models. Given the comparatively large bias and MSE associated with the placement parameter μ, it is not at all surprising that the associated estimates of σ_u^2 are much more heavily biased than those of σ_v^2. That is, if the mean cannot be reliably estimated from the data at hand, there should be no reason to expect the variance about that mean to be well estimated. Overall, the performance of the normal – truncated normal model appears to lend strong credence to the possibility of an identification problem, as suggested by Ritter and Simar

(1997a).

Though much less stark, certain differences in empirical performance also appear between the normal – half normal and normal – exponential model. Considering first the effects of sample size, it is apparent that the normal – exponential model is somewhat less sensitive than the normal – half normal specification. Except for when λ is large (4.743), the normal – exponential model consistently produces smaller bias and MSE in the estimation of σ_u^2, with the advantage over the normal – half normal model increasing as either sample size or λ grows small. Furthermore, accuracy in the estimation of σ_v^2 surpasses that of the normal – half normal model (in MSE terms) in all but one case[14]. In terms of bias, the estimation of σ_v^2 remains fairly comparable throughout, although the normal – exponential model appears to outperform the normal – half normal when both sample size and λ are small.

If we next consider the Jondrow, Lovell, Materov and Schmidt (JLMS) and Battese and Coelli (BC) point estimators of firm level technical efficiency, some further patterns are readily identifiable. First and foremost, it appears that the theoretically superior BC estimator does not produce estimates that are markedly different than those of the JLMS estimator. Neither estimator appears to be adversely affected by sample size, at least insofar as MSE is concerned. As was the case with the parameter estimates, however, the effects of λ are once again pronounced. Across all three models and regardless of sample size, both the BC and JLMS estimators show a decline in MSE, most notably when λ becomes large (4.473). It also appears that when sample size grows

---

[14]This occurs when $\lambda = 4.743$ and $n = 350$, with a difference in MSE of only $1\text{x}10^{-6}$.

small, any advantage of the BC estimator depreciates as λ grows larger.

It is interesting to note that irrespective of the comparatively poor estimation of σ_v^2, σ_u^2 and μ in the normal – truncated normal case, when both sample size and λ are large, the point estimators exhibit a smaller MSE than those of either the normal – half normal or normal – exponential cases. For smaller samples and/or smaller values of λ, however, the decline in the performance of these same point estimators is without parallel.

Comparing the point estimators between the normal – half normal and normal – exponential models, it appears that without exception both the BC and JLMS estimators are most accurate in the normal – exponential case, with the disparity increasing in all sample sizes as λ grows smaller. As previously noted, the normal – exponential model provides sharper estimation of σ_u^2, an artifact of which appears to be the increased accuracy of the subsequent point estimators.

The interval estimators appear to follow similar trends to the point estimators in regards to sample size. As would be expected, all three models exhibit a decrease in average coverage rate as the sample size falls, though this is clearly the most prevalent in the case of the normal – truncated normal specification. Not surprisingly, the empirical coverage rate for the normal – exponential model is superior to that of the normal – half normal model in all cases where sample size is larger than 50.

In contrast to the point estimators, however, the effect of λ varies according to the model under consideration. In the case of the normal – half normal specification, for example, it appears that the median value of $\lambda = 2$ yields the greatest empirical coverage rate, regardless of sample size. By contrast, the normal – exponential model achieves the greatest coverage rate, again regardless of sample size, when $\lambda = 1.5$. Finally, in the case of the normal – truncated normal model, no consistently optimal value of λ appears.

While the normal – half normal and normal – exponential models out perform the normal – truncated normal in terms of average coverage, it also seems abundantly clear that the normal – exponential model has an advantage, particularly as sample size becomes small. This is true for all but one case, where λ is large and the sample is extremely small ($n = 50$).

Finally, in considering Coelli's normally distributed test statistic for the presence of skewness, we find comparable trends to the interval estimators. The effect of sample size, for example, appears to be disparate across models. In particular, the normal – exponential model appears to be markedly more robust to reductions in sample size than either of the other models. Indeed, in every case examined, the p-values from the normal – exponential model easily carry the largest significance. Given this disparity, it would appear that skewness is most easily identified when the distribution of inefficiency is exponential.

The effect of λ contrasts with the pattern observed for the interval estimator: the significance of the test statistic increased in every model as λ grew larger. This result should by no means be surprising. The greater the presence of inefficiency, the greater must be the skewness in the distribution of ε. For this reason, and regardless of sample size, we should only expect to obtain a more significant test statistic as λ increases.

5.2 Simulation Results: Experiment Two

Experiment two considers the case where the data generating process differs from the assumptions of the model being estimated. Specifically we test the normal – half normal model when the data generating process is normal – exponential. Unsuccessful attempts were made to consider the converse, where the data generating process is normal – half normal and the normal – exponential model is estimated. In no case was consistent convergence attained. As a consequence, only results for the former case are provided.

As can be seen in tables 6 – 10, accuracy in parameter estimation falls considerably when the true data generating process differs largely from the assumptions of the normal – half normal model. In all but one case[15], and in samples of every size, both bias and MSE are markedly larger in the case of the misspecified model. Also, as would be expected, both bias and MSE for σ_u^2 are significantly worse than their σ_v^2 counter parts. The effect of increasing λ in the misspecified case is no different than when the model is correctly specified, with both bias and MSE improving as λ grows larger. One possibility for this may be the particular misspecification considered. That is, for the chosen values of λ, the normal – half normal and normal – exponential distributions for ε do not vary drastically in their general shape.

Poorer estimation of the critical σ_v^2 and σ_u^2 parameters carries over directly into the accuracy of the point estimates, as can be readily seen by considering the differences in average MSE for both the BC and JLMS estimators. When the model is misspecified, the effect of a reduction in sample size is pronounced, but also is not what would be

[15]This occurs at $n = 50$ and $\lambda = 4.743$.

expected. It is interesting note, for example, that in sample sizes of 50 and 100, the misspecified model produced more accurate point estimates (both BC and JLMS) than the correctly specified counterpart. In fact, it appears the misspecified model suffers considerably less from reductions in sample size than does the correctly specified model. The effect of λ coincides with what was observed in the case of the parameter estimates. That is, average MSE falls in both the misspecified and correctly specified models when λ grows large. It is also worth noting that regardless of both λ and the size of the sample, the BC estimator consistently outperforms the JLMS estimator in the misspecified model.

Interval estimation when the model is misspecified share both similarities and contrasts with point estimation. Turning first to the effects of sample size, we note that, like the point estimators, sample size increases the accuracy of the interval estimators, even when the model is misspecified. As expected, the interval estimators perform significantly worse when the model is misspecified, though not terribly so for large enough samples.

The effects of λ, however, strongly contrast what was observed with the point estimators. As λ grows large, interval estimation in fact suffers a consistent decrease in average coverage. This is perhaps counter intuitive, considering that both point estimators achieved smaller MSE as λ was increased. One possible explanation may be that the interval estimators rely more heavily upon the distributional characteristics of the underlying data generating process than do the conditional mean based point estimators.

Lambda = 1.5

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.003295	-0.030643	-0.004513	-0.001391	-0.044399	1.070880	-1.772627
MSE	0.007798	0.059259	0.003546	0.015481	0.027019	7.056824	27.081316
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.040070	0.040691	0.036274	0.037726	0.049607	0.049722	
Interval Est.							
Avg. Coverage	0.868100		0.892900		0.747860		
Coelli							
Avg. P-value	0.056010		0.000000		0.027620		

Lambda = 2

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	-0.001521	-0.007176	-0.001956	-0.002902	-0.020103	0.273020	-0.388766
MSE	0.001977	0.019867	0.001132	0.009196	0.005569	0.250177	0.957884
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.027427	0.027787	0.025768	0.026350	0.029348	0.029912	
Interval Est.							
Avg. Coverage	0.882580		0.892880		0.802420		
Coelli							
Avg. P-value	0.005148		0.000000		0.001633		

Lambda = 4.743

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	-0.001600	0.001768	-0.000487	-0.003601	-0.005262	0.212337	-0.323867
MSE	0.000132	0.006293	0.000113	0.007613	0.000279	0.084639	0.176974
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.009366	0.009399	0.008631	0.008659	0.008326	0.008440	
Interval Est.							
Avg. Coverage	0.873720		0.883380		0.805180		
Coelli							
Avg. P-value	0.000014		0.000000		0.000367		
Replications	100						
Observations	500						
Alpha	.10						

Table 1

Lambda = 1.5

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	-0.000321	-0.053137	-0.002065	-0.003938	-0.074796	1.043683	-1.737839
MSE	0.009147	0.085542	0.005351	0.021336	0.040718	6.638323	28.224029
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.042727	0.043031	0.036407	0.037809	0.054093	0.053824	
Interval Est.							
Avg. Coverage	0.851629		0.892657		0.689257		
Coelli							
Avg. P-value	0.114601		0.000000		0.059562		

Lambda = 2

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	-0.003286	-0.022910	-0.001626	-0.002815	-0.036107	0.564264	-0.938219
MSE	0.002488	0.029793	0.001745	0.012789	0.011457	3.474378	14.336280
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.028205	0.028320	0.025932	0.026513	0.034165	0.034092	
Interval Est.							
Avg. Coverage	0.873543		0.890771		0.722314		
Coelli							
Avg. P-value	0.019248		0.000000		0.010639		

Lambda = 4.743

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	-0.002042	-0.007144	-0.001455	-0.002374	-0.009526	0.207049	-0.300705
MSE	0.000174	0.009630	0.000175	0.010289	0.000462	0.101021	0.188440
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.009604	0.009588	0.008889	0.008911	0.008844	0.008930	
Interval Est.							
Avg. Coverage	0.862657		0.868314		0.725943		
Coelli							
Avg. P-value	0.000001		0.000000		0.000130		

Replications	100
Observations	350
Alpha	.10

Table 2

Lambda = 1.5

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.010299	-0.083529	0.005967	-0.025519	-0.093051	1.591138	-2.571055
MSE	0.016454	0.168576	0.010713	0.041619	0.055046	14.915478	50.257870
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.050935	0.050659	0.037287	0.038199	0.061773	0.060661	
Interval Est.							
Avg. Coverage	0.805850		0.884100		0.621400		
Coelli							
Avg. P-value	0.247615		0.000285		0.146159		

Lambda = 2

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.003521	-0.043974	0.001317	-0.013669	-0.045128	0.953606	-1.603694
MSE	0.004905	0.061848	0.003106	0.024973	0.016054	8.050329	29.041472
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.030642	0.030260	0.026430	0.026820	0.038181	0.037277	
Interval Est.							
Avg. Coverage	0.849900		0.882950		0.651500		
Coelli							
Avg. z-stat	-2.382271		-7.379265		-2.453056		
Avg. P-value	0.088900		0.000009		0.058885		

Lambda = 4.743

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	-0.001172	-0.010269	-0.001776	-0.002810	-0.010783	0.251549	-0.353897
MSE	0.000367	0.018550	0.000277	0.026038	0.000792	0.205416	0.399631
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.010109	0.010011	0.009261	0.009234	0.010378	0.010221	
Interval Est.							
Avg. Coverage	0.829400		0.845550		0.605450		
Coelli							
Avg. P-value	0.001141		0.000000		0.003444		

Replications	100
Observations	200
Alpha	.10

Table 3

Lambda = 1.5

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.032125	-0.161183	0.002385	-0.020144	-0.099555	1.998743	-2.988261
MSE	0.034983	0.377693	0.020373	0.109736	0.089868	25.940308	62.446369
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.090751	0.090794	0.039148	0.039424	0.076670	0.074649	
Interval Est.							
Avg. Coverage	0.629800		0.865800		0.465500		
Coelli							
Avg. P-value	0.394866		0.013589		0.334742		

Lambda = 2

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.021508	-0.106196	-0.002898	-0.006744	-0.042599	1.293848	-2.192180
MSE	0.013193	0.168422	0.006311	0.070388	0.028439	11.818552	35.644762
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.050777	0.049964	0.028233	0.028234	0.049370	0.047159	
Interval Est.							
Avg. Coverage	0.731700		0.857700		0.493400		
Coelli							
Avg. P-value	0.277811		0.001914		0.216339		

Lambda = 4.743

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.002908	-0.029218	-0.003526	0.001581	-0.008794	0.710070	-1.160310
MSE	0.001740	0.045966	0.000666	0.059001	0.003734	4.388032	14.943988
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.014404	0.014086	0.010844	0.010671	0.016267	0.015510	
Interval Est.							
Avg. Coverage	0.705300		0.744300		0.352000		
Coelli							
Avg. P-value	0.022773		0.000021		0.051265		
Replications	100						
Observations	100						
Alpha	.10						

Table 4

Lambda = 1.5

Model	**Half Normal**		**Exponential***		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.005050	-0.145652			-0.139149	1.383436	-1.812051
MSE	0.052879	0.576358			0.125973	19.831043	36.950970
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.114471	0.114484			0.084733	0.083424	
Interval Est.							
Avg. Coverage	0.496600				0.290400		
Coelli							
Avg. P-value	0.469640		.162432		0.483642		

Lambda = 2

Model	**Half Normal**		**Exponential***		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.010491	-0.114913			-0.069937	1.232886	-1.737728
MSE	0.021611	0.259718			0.039955	13.753374	30.819503
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.077878	0.077408			0.050875	0.049332	
Interval Est.							
Avg. Coverage	0.545600				0.313600		
Coelli							
Avg. P-value	0.398434		.084458		0.391076		

Lambda = 4.743

Model	**Half Normal**		**Exponential**		**Truncated Normal**		
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2	σ_v^2	σ_u^2	μ
Bias	0.005419	-0.064481	-0.012977	0.047640	-0.011838	0.913591	-1.392415
MSE	0.004348	0.094564	0.002278	0.144414	0.005221	5.852274	13.105606
Point Est.	BC	JLMS	BC	JLMS	BC	JLMS	
Avg. MSE	0.024085	0.023381	0.015992	0.015464	0.018963	0.017914	
Interval Est.							
Avg. Coverage	0.438200		0.370600		0.177000		
Coelli							
Avg. P-value	0.157767		0.015004		0.222236		
Replications	100						
Observations	50						
Alpha	.10						

Table 5 [16]

[16] * Failed to converge.

Lambda = 1.5

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.003295	-0.030643	-0.141845	1.794829
MSE	0.007798	0.059259	0.022959	3.311106
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.040070	0.040691	0.053531	0.061732
Interval Est.				
Avg. Coverage	0.868100		0.747260	
Coelli				
Avg. P-value	0.056010		0.000000	

Lambda = 2

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	-0.001521	-0.007176	-0.080372	1.409451
MSE	0.001977	0.019867	0.007196	2.034841
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.027427	0.027787	0.038968	0.043163
Interval Est.				
Avg. Coverage	0.882580		0.718000	
Coelli				
Avg. P-value	0.005148		0.000000	

Lambda = 4.743

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	-0.001600	0.001768	-0.020386	1.192574
MSE	0.000132	0.006293	0.000458	1.456488
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.009366	0.009399	0.013497	0.014073
Interval Est.				
Avg. Coverage	0.873720		0.631120	
Coelli				
Avg. P-value	0.000014		0.000000	
Model	Half Normal			
Replications	100			
Observations	500			
Alpha	.10			

Table 6

Lambda = 1.5

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	-0.000321	-0.053137	-0.142441	1.800177
MSE	0.009147	0.085542	0.024968	3.380740
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.042727	0.043031	0.053918	0.062048
Interval Est.				
Avg. Coverage	0.851629		0.739429	
Coelli				
Avg. P-value	0.114601		0.000000	

Lambda = 2

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	-0.003286	-0.022910	-0.082056	1.416797
MSE	0.002488	0.029793	0.007949	2.084301
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.028205	0.028320	0.039492	0.043620
Interval Est.				
Avg. Coverage	0.873543		0.706857	
Coelli				
Avg. P-value	0.019248		0.000000	

Lambda = 4.743

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	-0.002042	-0.007144	-0.021193	1.192588
MSE	0.000174	0.009630	0.000524	1.480634
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.009604	0.009588	0.013805	0.0143471
Interval Est.				
Avg. Coverage	0.862657		0.605200	
Coelli				
Avg. P-value	0.000001		0.000000	

Model	Half Normal
Replications	100
Observations	350
Alpha	.10

Table 7

Lambda = 1.5

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.010299	-0.083529	-0.135608	1.741805
MSE	0.016454	0.168576	0.025721	3.289359
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.050935	0.050659	0.053441	0.061409
Interval Est.				
Avg. Coverage	0.805850		0.742900	
Coelli				
Avg. P-value	0.247615		0.000285	

Lambda = 2

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.003521	-0.043974	-0.079449	1.383200
MSE	0.004905	0.061848	0.008097	2.069113
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.030642	0.030260	0.039238	0.043295
Interval Est.				
Avg. Coverage	0.849900		0.707900	
Coelli				
Avg. P-value	0.088900		0.000009	

Lambda = 4.743

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	-0.001172	-0.010269	-0.021837	1.183034
MSE	0.000367	0.018550	0.000598	1.538856
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.010109	0.010011	0.014176	0.014672
Interval Est.				
Avg. Coverage	0.829400		0.572750	
Coelli				
Avg. P-value	0.001141		0.000000	

Model	Half Normal
Replications	100
Observations	200
Alpha	.10

Table 8

Lambda = 1.5

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.032125	-0.161183	-0.144289	1.751367
MSE	0.034983	0.377693	0.037597	3.744582
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.090751	0.090794	0.054811	0.061835
Interval Est.				
Avg. Coverage	0.629800		0.705800	
Coelli				
Avg. P-value	0.394866		0.013589	

Lambda = 2

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.021508	-0.106196	-0.085088	1.386383
MSE	0.013193	0.168422	0.011593	2.311236
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.050777	0.049964	0.040614	0.044129
Interval Est.				
Avg. Coverage	0.731700		0.661300	
Coelli				
Avg. P-value	0.277811		0.001914	

Lambda = 4.743

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.002908	-0.029218	-0.024362	1.160574
MSE	0.001740	0.045966	0.000930	1.626699
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.014404	0.014086	0.015427	0.015777
Interval Est.				
Avg. Coverage	0.705300		0.407100	
Coelli				
Avg. P-value	0.022773		0.000021	

Model	Half Normal
Replications	100
Observations	100
Alpha	.10

Table 9

Lambda = 1.5

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.005050	-0.145652	-0.147789	1.670105
MSE	0.052879	0.576358	0.065001	4.239038
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.114471	0.114484	0.063378	0.068003
Interval Est.				
Avg. Coverage	0.496600		0.584200	
Coelli				
Avg. P-value	0.469640		0.162432	

Lambda = 2

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.010491	-0.114913	-0.096397	1.341460
MSE	0.021611	0.259718	0.020782	2.521440
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.077878	0.077408	0.046945	0.049163
Interval Est.				
Avg. Coverage	0.545600		0.478600	
Coelli				
Avg. P-value	0.398434		0.084458	

Lambda = 4.743

Inefficiency	**Half Normal**		**Exponential**	
Param.	σ_v^2	σ_u^2	σ_v^2	σ_u^2
Bias	0.005419	-0.064481	-0.030536	1.046162
MSE	0.004348	0.094564	0.001394	1.511050
Point Est.	BC	JLMS	BC	JLMS
Avg. MSE	0.024085	0.023381	0.017914	0.018056
Interval Est.				
Avg. Coverage	0.438200		0.163200	
Coelli				
Avg. P-value	0.157767		0.016374	

Model	Half Normal
Replications	100
Observations	50
Alpha	.10

Table 10

Chapter 6

Conclusion

For the purposes of this study simulated data were used to measure the empirical accuracy and sensitivity of both point and interval estimators under a variety of data generating processes and sample sizes. Also considered was the effect of these changes in the data generating process and reductions in sample size on the robustness of Coelli's (1995) asymptotic test for skewness in the OLS residuals.

Several conclusions can be drawn from the first of the two experiments, when the assumptions about the data generating process hold. First, regardless of sample size and the value of λ, the BC and JLMS point estimators do not produce largely differing estimates of firm level technical efficiency. Instead, it appears that the BC estimator provides only slightly more accuracy than the JLMS estimator. Furthermore, the advantage of the BC estimator appears to vanish when λ is sufficiently large and sample size becomes small. This is most notable in the normal – half normal case, where the JLMS estimator can be observed to out perform the BC estimator for both median and large values of λ with sample sizes of 200 or less.

The second conclusion to be drawn from the first experiment is that in all models, empirical coverage rates appeared to fall somewhat short of theoretical expectations[17]. This is readily apparent in the case of the normal – truncated normal model, but can also be seen in the normal – half normal and normal – exponential models as sample size falls. Also of interest is that a large value of λ, that is a greater presence of inefficiency in the

[17] That is, in no case did the interval estimates obtain an average coverage rate that met or exceeded the confidence level of 90%.

data, does not always lead to greater accuracy of the interval estimators. It was instead found that small λ generated the most accurate intervals in the normal – exponential case, and that median values of λ provide the sharpest interval estimates in the normal – half normal case. No consistently optimal value of λ appeared to exist in the case of the normal – truncated normal model.

In considering the normal – truncated normal model as a whole, we find strong evidence that an identification problem may exist, particularly with regards to the estimation of μ. It was clearly seen that poor estimation of this critical placement parameter greatly biased estimates of the distribution's spread, thereby leading to poorer point and interval estimation in all cases. Additional experimentation revealed that when sample size grew to approximately 2,000 observations, the normal – truncated normal model could be estimated as reliably and accurately as the normal – half normal and normal – exponential model. It would thus appear that the additional mean parameter μ does in fact require a large number of observations to be properly estimated. This is precisely the kind of identification problem hypothesized by Ritter and Simar (1997a).

In regard to Coelli's (1995) test statistic for the presence of inefficiency, we find greater robustness to changes in sample size than might have otherwise been expected, particularly given the asymptotic nature of the test statistic itself. It also appears that the test statistic is far more robust to changes in sample size in the normal – exponential model than the normal – half normal. In fact, for large enough values of λ, the test statistic in the normal – exponential model correctly identified[18] the presence of skewness

[18]That is, the p-value is less than .05.

with a sample size as small as 50. In general, however, the simulation results indicate that where λ is large, the test statistic remains reliable in sample sizes as small as 100. One might summarize these results as indicating that while large inefficiencies can be readily identified regardless of sample size, small inefficiencies require larger samples to be deemed significant at conventional levels.

The second experiment conducted concerned the implications of misspecification, particularly where the data generating process is normal – exponential and a normal – half normal model is applied. The results from this experiment lead to several interesting conclusions regarding the use of the normal – half normal model, and its normal – exponential counter part. Insofar as the point estimators are concerned, it appears that in every case the BC estimator is superior to the JLMS estimator, despite the model being misspecified. Furthermore, an increase in the sample size improves the accuracy of even the misspecified model's point estimates. Improvements in accuracy are also observed when λ grows larger.

The results from the interval estimators are somewhat more surprising. In terms of sample size the results are identical to the point estimators, in that an increase in sample size leads to increased accuracy (in terms of average coverage rate) of the interval estimators. The effect of λ however, is precisely the opposite of what was observed with the point estimators. That is, the average coverage rate for the misspecified model appears to decrease notably as λ grows large. This can be observed for all sample sizes, but is clearly the most prominent when the sample size falls. This same general behavior was also observed when the model was correctly specified, and thus may be an artifact of the particular form of misspecification applied in the experiment.

On the whole, the results of this analysis seem to suggest that the asymptotic properties underlying the point estimators tend to hold even for somewhat small samples, provided that significant inefficiency exists. In the absence of at least median values of λ or samples larger than 200, however, caution indeed appears warranted. This is particularly true when the misspecification of the underlying distribution is suspected. While the remarks on sample size apply equally to the interval estimators, those concerning λ do not. The optimal value of λ appears to be dependent upon the underlying distribution of the inefficiency, and in general a large value of λ does not lead to higher empirical coverage rates. This somewhat paradoxical relationship between point estimation, where large λ is optimal, and interval estimation where it is not, may warrant future research.

References

Aigner, Dennis J. and S.F. Chu. “On Estimating the Industry Production Function.” *American Economic Review* 58(September 1968): 826-839.

Aigner, Dennis J., C.A. Knox Lovell, and Peter Schmidt. “Formulation and Estimation of Stochastic Frontier Production Function Models.” *Journal of Econometrics* 6(1977): 21-37.

Battese, G.E. and T.J. Coelli. “Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data.” *Journal of Econometrics* 38(1988): 387-399.

Bera, Anil K. and Subhash C. Sharma. “Estimating Production Uncertainty in Stochastic Frontier Production Function Models.” *Journal of Productivity Analysis* 12(November 1999): 187-210.

Bera, Anil K. and Naresh C. Mallick. “Information Matrix Tests for the Composed Error Frontier Model.” *Advances on Methodological and Applied Aspects of Probability and Statistics*. N. Balakrishnan, ed. New Zealand: Macmillan, 2002.

Coelli, T. “Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis.” *Journal of Productivity Analysis* 6(1995): 247-68.

Farrel, M.J. “The Measurement of Productive Efficiency.” *Journal of the Royal Statistical Society* A120(1957): 253-281.

Greene, W.H. “Maximum Likelihood Estimation of Econometric Frontier Functions.” *Journal of Econometrics* 13(May 1980): 27-56.

Greene, W.H. “A Gamma-Distributed Stochastic Frontier Model.” *Journal of Econometrics* 46 (October 1990): 141-164.

Horrace, William C. and Peter Schmidt. “Confidence Statements for Efficiency Estimates from Stochastic Frontier Models.” *Journal of Productivity Analysis* 7(July 1996): 257-282.

Jondrow, James, C.A. Knox Lovell, Ivan S. Materov, and Peter Schmidt. “On the Estimation of Technical Efficiency in the Stochastic Frontier Production Function Model.” *Journal of Econometrics* 19(1982): 233-238.

Lee, Y.H. and W.G. Tyler. “The Stochastic Frontier Production Function and Average Efficiency.” *Journal of Econometrics* 7(June 1978): 385-389.

Meeusen, Wim, and Julien van Den Broeck. “Efficiency Estimation from Cobb-Douglas Production Functions with Composed Error.” *International Economic Review* 18(June 1977): 435-444.

Olson, J.A., Peter Schmidt and D.M. Waldman. “A Monte Carlo Study of Estimators of Stochastic Frontier Production Functions.” *Journal of Econometrics* 13(May 1980): 67-82.

Ritter, C. and L. Simar. “Pitfalls of Normal-Gamma Stochastic Frontier Models.” *Journal of Productivity Analysis* 8(May 1997): 167-182.

Ritter, C. and L. Simar. “Another Look at the American Electrical Utility Data.” Working paper, CORE and Institut de Statistique, 1997.

Seits, W.D. “Productive Efficiency in the Steam-Electric Generating Industry.” *Journal of Political Economy* 79(July 1971): 878-886.

Stevenson, Rodney E. “Likelihood Functions for Generalized Stochastic Frontier Estimation.” *Journal of Econometrics* 13(1980): 57-66.

Subal C. Kumbhakar and C.A. Knox Lovell. *Stochastic Frontier Analysis*. United Kingdom: Cambridge University Press, 2000.

Timmer, C.P. “Using a Probilistic Frontier Production Function to Measure Technical Efficiency.” *Journal of Political Economy* 79(July 1971): 776-794.