This is to certify that the

dissertation entitled

A Study in Estimation and Inference on Firm Efficiency

presented by

Yangseon Kim

has been accepted towards fulfillment
of the requirements for

Ph.D degree in Economics

Peter Schmidt
Major professor

Date 5-12-1999

A STUDY IN ESTIMATION AND INFERENCE ON FIRM EFFICIENCY

By

Yangseon Kim

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

1999

ABSTRACT

A STUDY IN ESTIMATION AND INFERENCE ON FIRM EFFICIENCY

By

Yangseon Kim

This thesis considers the problem of interval estimation of technical efficiency levels in stochastic frontier models with panel data. First we consider a large number of classical and Bayesian procedures to estimate technical efficiency levels of firms and to construct confidence intervals for these efficiency levels. We then apply these methods to three data sets with different characteristics. The fixed effects models generally perform poorly: there is a large payoff to distributional assumptions for efficiencies. We do not find much difference between Bayesian and classical methods if we match methods that depend on comparable assumptions.

This thesis also provides simulation evidence on the accuracy of inferences on technical efficiency levels. The simulation evidence suggests that MCB and MargCB are very conservative. Inference based on bootstrapping is not very reliable in general, but it is reasonably reliable when T is large relative to N and when the variance of the inefficiency term is large relative to the variance of the noise. Using the bias-adjusted bootstrap helps but does not solve the problem completely. The method which uses a distributional assumption in estimating inefficiencies, performs surprisingly poorly when N is small. However, there appears to be a large gain in precision and reliability of inference from making a distributional assumption when N is large relative to T.

ACKNOWLEDGMENTS

The completion of this study owes much to the support and encouragement of many people who deserve recognition for their efforts. I am especially indebted to my dissertation adviser, Peter Schmidt. I am deeply appreciative of the guidance and encouragement which I received from him over the period of this study. This work would not have appeared in its present form without him. I would also like to thank other members of the thesis committee, Jeff Wooldridge, Robert de Jong and Richard Baillie for their time and effort and for their insightful comments.

Many of my fellow graduate students have also contributed to the completion of this study through their encouragement and friendship. My special thanks go out to Heather Bednarek, Hailong Qian, and Yi-Yi Chen. A very warm thank you is extended to Ann Feldman and Zeynep Altinsel who have stayed with me through some tough times. Their kind advise and warm friendship will always be appreciated.

Finally, I reserve special thank to my family. The person who deserves the most recognition is my husband, Sang Hyop Lee. His love and support are greatly appreciated. I would also like to thank my lovely two sons, Jihoon and Sehoon, for their patience and love. They tried to understand me even though they frequently missed my attention. I would never have arrived at this point without their smile. They deserve more than half the credit that I may have achieved.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 2

CHAPTER 3

CHAPTER 4

Chapter 1

INTRODUCTION

This thesis consists of three essays on the problem of estimating the technical efficiency of firms, using panel data. This chapter will define some basic notation and give an overview of the thesis.

1. The Basic Model

We begin with the basic panel data stochastic frontier model of Pitt and Lee (1981) and Schmidt and Sickles (1984):

$$(1.1) \qquad y_{it} = \alpha + x_{it}'\beta + v_{it} - u_i, \quad i= 1,.....N, \quad t=1,.....,T$$

Here i indexes firms or productive units and t indexes time periods. y_{it} is the scalar dependent variable representing the logarithm of output for the i^{th} firm in period t, α is a scalar intercept, x_{it} is a Kx1 vector of functions of inputs (e.g., in logarithms for the Cobb-Douglas specification), β is a Kx1 vector of coefficients and v_{it} is an i.i.d. error term with zero mean and finite variance. The u_i satisfy $u_i \geq 0$, and $u_i > 0$ is an indication of technical inefficiency. Note that u_i is time-invariant. For a logarithmic specification such as this the technical efficiency of the i^{th} firm is defined as $r_i = \exp(-u_i)$, so technical inefficiency is $1-r_i$. For small values of u_i, u_i is approximately equal to $1- \exp(-u_i) = 1-r_i$, so that u_i itself is sometimes used as a measure of technical inefficiency.

Now define $\alpha_i = \alpha - u_i$. With this definition, (1.1) becomes the standard panel data model with time-invariant individual effects :

(1.2) $$y_{it} = \alpha_i + x_{it}'\beta + v_{it}$$

Obviously we have $\alpha_i \leq \alpha$ and $u_i = \alpha - \alpha_i$. As before, technical efficiency is $r_i = \exp(-u_i)$.

The previous discussion regards zero as the minimal possible value of u_i, and α as the maximal possible value of α_i, over any possible sample; that is, essentially, as $N \to \infty$. For some purposes, and especially when N is not large, it is also useful to consider the following representation. We write the intercepts α_i, in ranked order as :

(1.3) $$\alpha_{(1)} \leq \alpha_{(2)} \ldots \leq \alpha_{(N)}$$

so that in particular (N) is the index of the firm with the largest value of α_i, among firms i=1,...,N. It is convenient to write the values of u_i in the opposite ranked order, as $u_{(N)} \leq \ldots \leq u_{(2)} \leq u_{(1)}$, so that $\alpha_{(i)} = \alpha - u_{(i)}$ for all i. Then obviously $\alpha_{(N)} = \alpha - u_{(N)}$, and firm (N) has the largest value of α_i or equivalently the smallest value of u_i among firms i=1,...,N. We will call this firm the best firm in the sample. In some methods we measure inefficiency relative to the *best* firm in the sample, and this corresponds to considering the relative efficiency measures:

(1.4) $$u_i^* = u_i - u_{(N)} = \alpha_{(N)} - \alpha_i, \quad r_i^* = \exp(-u_i^*).$$

2. Outline of the Thesis

Chapter 2 is primarily empirical. It first provides a survey of a large number of classical and Bayesian methods that have been proposed to estimate technical efficiency

levels and to perform inference on these levels. Classical procedures include multiple comparisons with the best (MCB), based on the fixed effects estimates; marginal comparisons with the best (MargCB); bootstrapping of the fixed effects estimates; and maximum likelihood given a distributional assumption for the inefficiency terms u_i. Bayesian procedures include a Bayesian version of the fixed effects model, and various Bayesian models with informative priors for efficiencies.

It then applies these techniques to the three previously-analyzed data sets, on Indonesian rice farms, Texas electric utilities, and Egyptian tileries, and compares the point estimates and confidence intervals for technical efficiency levels. The fixed effects models generally perform poorly; there is a large payoff to distributional assumptions for efficiencies. There is not much difference between Bayesian and classical procedures, in the sense that classical MLE based on a distributional assumption for efficiencies gives results that are rather similar to a Bayesian analysis with the corresponding prior.

Chapter 3 provides simulation evidence on the accuracy of inferences on technical efficiency levels. This is useful because the finite sample properties of the efficiency measurement methods discussed in Chapter 2 are basically unknown. It is known that MCB and MargCB confidence intervals are considerably wider than intervals based on bootstrapping the fixed effects estimates or based on a distributional assumption, but it is not known how conservative MCB and MargCB are, or whether the bootstrapping or distribution-based estimates are reliable. The simulation evidence suggests that MCB and MargCB are *very* conservative. Inference based on bootstrapping is not very reliable in general, but it is reasonably reliable when T is large relative to N and when the variance of the inefficiency term u_i is large relative to variance of the noise v_{it}. The main problem

with bootstrapping is finite sample bias due to the "max" operation in defining the frontier, and using the bias-adjusted bootstrap helps but does not solve the problem completely. There appears to be a large gain in precision and reliability of inference from making a distributional assumption. This is so especially when N is large relative to T.

Chapter 4 provides a rigorous derivation of Marginal Comparisons with the Best (or MargCB) which was introduced in Chapter 2. MargCB addresses the following problem. Given estimates of the basic model, we wish to say which populations might be best, and to construct confidence intervals for the differences $u_i^* = \theta_{(N)} - \theta_i$, which measure the amount by which a given population differs from the best. An existing technique called Multiple Comparisons with the Best (or MCB) also addresses this problem. MCB constructs a set S of possibly best populations, and a set of intervals (L_i, U_i), such that:

$$(1.5) \qquad P[(N) \in S \text{ and } L_i \le u_i^* \le U_i \text{ for all } i] \ge (1-c),$$

where 1-c is a chosen confidence level (e.g., 0.90). Thus with a given confidence level we have a set of populations that includes the best, and joint confidence intervals for all differences from the best. These intervals are often rather wide, in part because of the joint nature of the statement (1.5). One way to make the intervals narrower is to make a marginal rather than joint statement, of the form :

$$(1.6) \qquad P[(N) \in S \text{ and } L_i \le u_i^* \le U_i \,] \ge (1-c),$$

where statement (1.6) holds for a single given value of i. The MargCB technique provides the set S and interval (L_i, U_i) such that (1.6) is true. This technique is then applied to the

same three data sets as were used in Chapter 2, and evidence on the width of the MCB and MargCB intervals is provided.

Chapter 2

AN EMPIRICAL COMPARISON OF BAYESIAN AND CLASSICAL APPROACHES TO INFERENCE ON EFFICIENCY LEVELS IN STOCHASTIC FRONTIER MODELS WITH PANEL DATA

1. Introduction

This chapter considers the problem of interval estimation of technical efficiency levels in stochastic frontier models with panel data. The phrase "interval estimation" indicates that we are interested not only in point estimates of the efficiency levels of the individual firms, but also in confidence intervals for the efficiency levels.

A number of different techniques have been proposed in the literature to address this problem. Given a distributional assumption for technical inefficiency, maximum likelihood estimation was proposed by Pitt and Lee (1981). Battese and Coelli (1988) showed how to construct point estimates of technical efficiency for each firm, and Horrace and Schmidt (1996) showed how to construct confidence intervals for these efficiency levels. Without a distributional assumption for technical inefficiency, Schmidt and Sickles (1984) proposed fixed effects estimation, and the point estimation problem for efficiency levels was discussed by Schmidt and Sickles (1984) and Park and Simar (1994). Simar (1992) and Hall, Härdle and Simar (1993) suggested using bootstrapping to conduct inference on the efficiency levels. Horrace and Schmidt (1996, 1999) constructed confidence intervals using the theory of multiple comparisons with the best, and we extend it to a univariate version. Bayesian methods have been suggested by Koop, Osiewalski and Steel (1997) and Osiewalski and Steel (1998). They propose a model with an uninformative prior for firm-specific intercepts that is intended to be similar to the

classical fixed effects model, and also models with informative priors, which are comparable to classical models that assume a distribution for inefficiency.

These models have been applied to various data sets, but there has been no systematic attempt to compare them all on a common data set. In this chapter, we apply these models to three previously-analyzed data sets and compare the results. The major emphasis is to try to understand the relationship between the assumptions underlying the various models and the empirical results. More specifically, we are interested in two types of questions. First, we wish to see how much is gained, in terms of tightness of the confidence intervals, by being willing to make a distributional assumption for technical inefficiency. This is phrased as a classical question, but from a Bayesian perspective the question is simply rephrased as seeing how much is gained, in terms of tightness of the posterior distribution, by imposing an informative prior. Second, we wish to compare the results from Bayesian and classical analyses, where we match as far as possible the strength of the assumptions underlying the analyses. We find large gains from distributional assumptions, and we do not find much difference between classical and Bayesian analyses that rely on assumptions of comparable strength.

2. Classical Statistical Procedures

In this section we will discuss classical statistical procedures for the estimation of the model presented in section 1 of the chapter 1. Bayesian procedures will be discussed in the next section. We will distinguish procedures that make an assumption about the distribution of u_i from those that do not.

2.1 Efficiency Measurement with a Distributional Assumption for Inefficiency

In this subsection we consider estimation under strong assumptions that are similar to those made in the cross-sectional case (Aigner, Lovell and Schmidt (1977)). We assume independence across firms (values of i). We assume that the explanatory variables x_{it} are strictly exogenous: $(x_{i1},...,x_{iT})$ is independent of $(v_{i1},...,v_{iT},u_i)$. We assume that the v_{it} are i.i.d. as $N(0,\sigma_v^2)$ and are independent of u_i. Finally, we assume a specific distribution for u_i. The distributions most commonly considered in the literature have been the half-normal and the exponential. Other suggestions include the truncated normal (Stevenson (1980)) and gamma (Greene (1990)) distributions. In this chapter we will use the exponential distribution, primarily because it is easiest to handle in the Bayesian framework, and we want to be able to make comparisons across Bayesian and classical approaches based on the same distribution for u. Thus we assume the following density for u: $f(u) = \phi^{-1}\exp(-u/\phi)$, with $E(u) = \phi$.

The preferred method of estimation is maximum likelihood estimation (MLE). The likelihood function has been derived by Aigner, Lovell and Schmidt (1977) for the cross-sectional case (T=1), and is easily extended to the case of panel data as follows:

$$\ln L = -\ln(\phi) - \frac{N(T-1)}{2}\ln(2\pi\sigma_v^2) - \frac{N}{2}\ln T - \frac{T}{2\sigma_v^2}\sum_{i=1}^{N}[\sum_{t=1}^{T}\frac{\varepsilon_{it}^2}{T} - (\bar{\varepsilon}_i + \frac{\sigma_v^2}{T\phi})] + \sum_{i=1}^{N}\ln[1 - \Phi(\frac{\sqrt{T}}{\sigma_v}\bar{\varepsilon}_i + \frac{\sigma_v}{\sqrt{T}\phi})] \tag{2.1}$$

where Φ denotes the cdf of the standard normal distribution, $\varepsilon_{it} = v_{it} - u_i = y_{it} - \alpha - x_{it}\beta$ and $\bar{\varepsilon}_i = T^{-1}\sum_{t=1}^{T}\varepsilon_{it}$. The likelihood function is maximized numerically to obtain the MLE

of the parameters (α, β, ϕ, σ_v^2). These estimates are consistent as $N\to\infty$ for fixed T. The implication of this is that N needs to be large for MLE to be appropriate, whereas large T is not required.

The model can also be regarded as the random effects model from the panel data literature, so that estimation by generalized least squares (GLS) is possible. GLS is consistent as $N\to\infty$ without the assumption of normality of the v_{it} and without an assumption of a specific distribution for the u_i. However, the model above differs from the standard random effects model because u_i does not have a mean of zero. Thus the GLS estimated intercept will be a consistent estimate of α-E(u) rather than of α. It is possible to adjust the intercept upward, by adding a consistent estimate of E(u), but this requires an assumption about the distribution of u. As a result we will consider only the MLE.

Estimation of the model yields residuals $\hat{\varepsilon}_i = y_{it} - \hat{\alpha} - x_{it}'\hat{\beta}$, which are naturally regarded as estimates of $\varepsilon_{it} = v_{it} - u_i$, whereas we are interested in estimating u_i itself. The usual solution to this problem, following Jondrow et al. (1982) and Battese and Coelli (1988), is to consider the distribution of u_i conditional on $\varepsilon_i = (\varepsilon_{i1},.....,\varepsilon_{iT})$ evaluated at $\hat{\varepsilon}_i$. Jondrow et al. give this distribution for the cross-sectional case (T=1), for the cases that u is half-normal or exponential, and suggest the point estimate $E(u_i \mid \varepsilon_i)$. Battese and Coelli (1988) give both $E(u_i \mid \varepsilon_i)$ and $E(r_i \mid \varepsilon_i)$ for the half-normal case with panel data ($T\geq 1$). Actually, when the v_{it} are i.i.d. normal, the distribution of u_i conditional on $\varepsilon_i = (\varepsilon_{i1},.....,\varepsilon_{iT})$ is the same as the distribution of u_i conditional on $\bar{\varepsilon}_i$, regardless of the distribution of u_i. Therefore the results of Jondrow et al. can be extended to the panel

data case just by replacing ε_i by $\bar{\varepsilon}_i$ and σ_v^2 by σ_v^2/T. From their Theorem 2, we find that the distribution of u_i conditional on ε_i is $N(-\mu_i, \sigma_v^2/T)$ truncated at zero, where $\mu_i = \bar{\varepsilon}_i + \sigma_v^2/(T\phi)$. We can therefore calculate confidence intervals using this distribution. This idea was suggested by Horrace and Schmidt (1996), who implemented it assuming the half-normal distribution for u, and the results in this chapter differ only because a different distribution (exponential) is assumed.

Conducting inference on u_i using the distribution of u_i conditional on ε_i is not suggested as an authentic Bayesian procedure, but it obviously has a Bayesian flavor. The main difference between this distribution and a Bayesian posterior distribution is that it relies on asymptotics to ignore the effects of parameter estimation, whereas the uncertainty due to parameter estimation will figure into the Bayesian posterior. We might expect this difference not to matter very much when N is large, however.

2.2 Estimation without a Distributional Assumption for Inefficiency: Fixed Effects

We now discuss estimation without a distributional assumption for the u_i. This subsection will give a brief review of the point estimation problem based on the fixed-effects estimates, and the next two subsections will consider different ways of constructing confidence intervals for inefficiency based on these estimates.

Fixed effects estimation refers to the estimation of the panel data regression model (1.2), treating the α_i as fixed parameters. Because the α_i are treated as parameters, we do not need to make any distributional assumption about the inefficiencies; nor do we need to assume that they are uncorrelated with the x_{it} or the v_{it}. We still assume strict exogeneity of the regressors x_{it} in the sense that $(x_{i1},\ldots,x_{iT})$ is independent of

$(v_{i1},\ldots,v_{iT})$. We also assume that the v_{it} have zero mean and constant variance σ_v^2, and are not autocorrelated. We do not need to assume a distribution for the v_{it}.

The fixed effects estimate $\hat{\beta}$, also called the *within* estimate, may calculated by regressing $(y_{it}-\bar{y}_i)$ on $(x_{it}-\bar{x}_i)$, or equivalently by regressing y_{it} on x_{it} and a set of N dummy variables for firms. We then obtain $\hat{\alpha}_i = \bar{y}_i - \bar{x}_i'\hat{\beta}$; or equivalently the $\hat{\alpha}_i$ are the estimated coefficients of the dummy variables. This leads to the following expression for $\hat{\alpha}_i$:

$$\hat{\alpha}_i = \alpha_i + \bar{v}_i - \bar{x}_i'(\hat{\beta}-\beta) \tag{2.2}$$

to which we will make reference later. The fixed effects estimate $\hat{\beta}$ is consistent as $NT\to\infty$ (i.e., as either N or T approaches infinity), and its variance is of order $[N(T-1)]^{-1}$. For a given firm (i), the estimated intercept $\hat{\alpha}_i$ is a consistent estimate of α_i as $T\to\infty$. Large T is needed for the term $\bar{v}_i$ in (2.2) to become negligible.

Schmidt and Sickles (1984) suggested the following estimates of technical inefficiency, based on the fixed effects estimates:

$$\hat{\alpha} = \max_{j=1,\ldots,N} \hat{\alpha}_j \; ; \quad \hat{u}_i^* = \hat{\alpha} - \hat{\alpha}_i, \; i=1,\ldots,N. \tag{2.3}$$

Since these estimates clearly measure inefficiency relative to the firm estimated to be the best in the sample, they are naturally viewed as estimates of $\alpha_{(N)}$ and u_i^*, that is, of relative rather than absolute inefficiency. For fixed N, $\hat{\alpha}$ is a consistent estimate of $\alpha_{(N)}$ and $\hat{u}_i^*$ is a consistent estimate of u_i^* as $T\to\infty$. However, it is important to note that in

finite samples (for small T) $\hat{\alpha}$ is likely to be biased upward, since $\hat{\alpha} \geq \hat{\alpha}_{(N)}$ and $E(\hat{\alpha}_{(N)}) = \alpha_{(N)}$, where $\hat{\alpha}_{(N)}$ is the estimated intercept for the *unknown* best firm. That is, the "max" operator in (2.3) induces upward bias, since the largest $\hat{\alpha}_i$ is more likely to contain positive estimation error than negative error. This bias is larger when N is larger and when the $\hat{\alpha}_i$ are estimated less precisely. The upward bias in $\hat{\alpha}$ induces an upward bias in the $\hat{u}_i^*$ and a downward bias in $\hat{r}_i^* = \exp(-\hat{u}_i^*)$; we underestimate efficiency because we overestimate the level of the frontier.

Schmidt and Sickles argued that $\hat{\alpha}$ and the $\hat{u}_i^*$ are consistent estimates of α and the u_i if both N and T approach infinity; that is, if both N and T are large, we can regard the $\hat{u}_i^*$ as estimates of absolute and not just relative inefficiency. The argument is simple. As $T \to \infty$, $\hat{\alpha}$ and the $\hat{u}_i^*$ are consistent estimates of $\alpha_{(N)}$ and the u_i^*, as noted above. As $N \to \infty$, $u_{(N)}$ should converge to zero, so that $\alpha_{(N)}$ should converge to α and the u_i^* should converge to the corresponding u_i. A more rigorous treatment of the asymptotics for this model is given by Park and Simar (1994), who show that, in addition to $N \to \infty$ and $T \to \infty$, we need to require $T^{-1/2} \ln N \to 0$ in order to ensure the consistency of $\hat{\alpha}$ as an estimate of α. This latter requirement limits the rate at which N can grow relative to T, in order to ensure that the upward bias induced by the "max" operation disappears asymptotically.

2.3 Multiple and Marginal Comparisons with the Best

Multiple comparisons with the best (MCB) is a statistical technique that yields confidence intervals for differences in parameter values between all populations and the

best population. Horrace and Schmidt (1996, 1998) have suggested its use to construct confidence intervals for the relative technical inefficiencies $u_i^* = \alpha_{(N)} - \alpha_i = u_i - u_{(N)}$ or $r_i^* = \exp(-u_i^*)$, which are indeed differences from the best.

Let $A = (\alpha_1,...,\alpha_N)$ be the vector of intercepts for the N firms in the panel data regression model (1.2). (It would be natural to refer to this vector as α, but that symbol has already been used for the intercept in model (1.1).) As before we denote $\max_{j=1,...,N} \alpha_j$ by $\alpha_{(N)}$. Then MCB constructs a set S of possibly best populations, and a set of intervals (L_i,U_i), such that:

(2.4) $$P[\,(N) \in S \textit{ and } L_i \le \alpha_{(N)} - \alpha_i \le U_i \text{ for all } i\,] \ge 1\text{-}c\,,$$

where 1-c is a chosen confidence level (e.g., 0.90). (Again, it would be natural to use α in place of c for the tail probability, but the symbol α has already been used.) Thus with a given confidence level we have a set of populations that includes the best, and joint confidence intervals for all differences from the best. MCB was developed by Hsu (1981, 1984) and Edwards and Hsu (1983). A general exposition can be found in Hochberg and Tamhane (1987), Hsu (1996) and Horrace and Schmidt (1999).

To perform MCB, we need an estimate $\hat{A}$, distributed as $N(A,\sigma^2 C)$ with C known, and where either σ^2 is known, or we have an estimate $\hat{\sigma}^2$, independent of $\hat{A}$, such that $\hat{\sigma}^2/\sigma^2$ is distributed as χ_v^2/v. In typical MCB applications to the efficiency measurement problem, $\hat{A}$ will come from the fixed-effects estimation of the panel data regression model (1.2), as discussed above, and there will be enough degrees of freedom

that we can effectively take σ^2 as known. Normality of $\hat{A}$ requires either that the errors v_{it} are normal or that T is large enough for a central limit theorem to hold for the expressions in (2.2) above.

Standard MCB proceeds under the further assumption that $C = kI_N$ with k known. This assumption is usually motivated by discussion of the "balanced one way model" (e.g., Hsu (1996), p. 43) in which we have independent observations y_{it} (i = 1,...,N, t = 1,...,T) distributed as $N(\alpha_i, \sigma^2)$. In this case k = 1/T. This is equivalent to the panel data regression model (1.2) if β were known, since then we have $(y_{it} - x'_{it}\beta) = \alpha_i + v_{it}$, which is the balanced one-way model. Since standard MCB would be applicable if β were known, it is reasonable to presume that standard MCB is a good approximation if β is estimated sufficiently precisely. Recall from the previous subsection that the variance of $\hat{\alpha}_i$ is of order T^{-1}, while the variance of $\hat{\beta}$ is of order $[N(T-1)]^{-1}$. Thus it may generally be the case that standard MCB is approximately applicable when N is large. This point is discussed in more detail in Horrace and Schmidt (1999).

Now define the following notation. Let E(½) be the (N-1)x(N-1) correlation matrix with all correlations equal to ½ (i.e., diagonal elements equal one, off-diagonal elements equal ½). Let z be a multivariate random variable distributed as student-t with dimension N-1, degrees of freedom v, and correlation matrix E(½). Define $d^*(c)$ as the c-level critical value of $\max_{i=1,...,N-1} |z_i|$; i.e., $P[\max_{i=1,...,N-1} |z_i| \leq d^*(c)] = 1-c$. Tabulations of $d^*(c)$ can be found in Hsu (1996) or Horrace (1998). Define $h(c) = d^*(c)(2k\hat{\sigma}^2)^{1/2}$, and define the set $S(c) = \{i \mid \hat{\alpha}_i \geq \max_{j=1,...,N} \hat{\alpha}_j - h(c)\}$. Define L_i and U_i as follows:

(2.5) $$L_i = \max[0, \min_{j \in S(c)} \hat{\alpha}_j - \hat{\alpha}_i - h(c)], \; U_i = \max[0, \max_{j \neq i} \hat{\alpha}_j - \hat{\alpha}_i + h(c)]$$

Then MCB provides the statement (2.4) above, with S = S(c).

As noted above, standard MCB requires that the variance matrix of $\hat{A} = (\hat{\alpha}_1, .., \hat{\alpha}_N)$ be proportional to an identity matrix; that is, the various $\hat{\alpha}_i$ are uncorrelated and have equal variance. *General* MCB allows this variance matrix to be of an arbitrary form. A discussion of general MCB can be found in Horrace and Schmidt (1999, section 3), and is too lengthy to include here. The empirical results in this paper are all obtained from general MCB, but standard MCB would have yielded similar results.

The MCB statement (2.4) is a multiple statement in the sense that the confidence intervals all hold jointly with (at least) the specified probability. We can also consider marginal (one at a time) confidence intervals, which are more directly comparable to the intervals provided by the other techniques we consider. We call this marginal comparisons with the best, which we will abbreviate as MargCB. There are *standard* and *general* versions of MargCB, where the standard version makes the same assumption about the variance matrix of $\hat{A}$ as does standard MCB. We will discuss standard MargCB, but the empirical results of this chapter use the general version. Let $t^*(c)$ be the two-sided c-level critical value of the (univariate) student-t distribution with v degrees of freedom; i.e., if z is distributed as student-t with v degrees of freedom, then $P[|z| \leq t^*(c)] = 1 - c$. Define $g(c) = t^*(c)(2k\hat{\sigma}^2)^{1/2}$. Define the set S(c) as above. Define L_i^m and U_i^m as follows:

(2.6) $$L_i^m = \max[0, \min_{j \in S(c)} \hat{\alpha}_j - \hat{\alpha}_i - g(c/2)],\ U_i^m = \max[0, \max_{j \neq i} \hat{\alpha}_j - \hat{\alpha}_i + g(c/2)].$$

Then the following is true:

(2.7) $$P[\ (N) \in S(c)\ \textit{and}\ L_i^m \leq \alpha_{(N)} - \alpha_i \leq U_i^m\ ] \geq 1\text{-}c.$$

A proof of this result in a more general setting will be given in Chapter 4.

Both MCB and MargCB are conservative procedures. The events given in statements (2.4) and (2.7) hold with a probability of *at least* 1-c, and the inequality occurs because of uncertainty about which firm is best. There will be considerable uncertainty about which firm is best when one or more of the α_i are nearly as large as $\alpha_{(N)}$ and when the $\hat{\alpha}_i$ have large sampling variance. In such cases the set S(c) will be large and the confidence intervals will be wide. The other techniques discussed in this chapter are not conservative and may be expected to yield narrower confidence intervals. However, those techniques rely on stronger assumptions and/or asymptotic theory and correspondingly may be more likely to yield inferences that are erroneous.

MCB is not designed as an asymptotic procedure. Indeed, the problem of comparing N populations is hard to conceptualize unless N is fixed. However, since econometricians often think in terms of asymptotics, the following comments may be helpful. First, as just noted, standard MCB may be approximately valid when N is large. Second, MCB assumes that the $\hat{\alpha}_i$ are normally distributed. This should be so if the errors v_{it} are normal *or* if T is large. Third, MCB establishes confidence intervals for the relative inefficiencies u_i^* or r_i^*, but if N is large these can also be regarded as confidence intervals for the absolute inefficiencies u_i or r_i.

2.4 Bootstrapping

We can use bootstrapping to construct confidence intervals for functions of the fixed effects estimates. The inefficiency measures $\hat{u}_i^*$ (as in equation (2.3)) and the efficiency measures $r_i^* = \exp(-\hat{u}_i^*)$ are functions of the fixed effects estimates and so bootstrapping can be used for inference on these measures.

Consider the general setting in which we have a parameter θ, and there is an estimate $\hat{\theta}$ based on a sample $z_1,...,z_n$ of i.i.d. random variables. The estimator $\hat{\theta}$ is assumed to be regular enough so that $n^{1/2}(\hat{\theta}-\theta)$ is asymptotically normal. The following bootstrap procedure will be repeated many times, say for b = 1,...,B where B is large. For iteration b, construct pseudo data $z_1^{(b)},...,z_n^{(b)}$ by sampling randomly with replacement from the original data $z_1,...,z_n$. From the pseudo data, construct the estimate $\hat{\theta}^{(b)}$. The basic result of the bootstrap is that, under fairly general circumstances, the asymptotic (large n) distribution of $n^{1/2}(\hat{\theta}^{(b)}-\hat{\theta})$ conditional on the sample is the same as the (unconditional) asymptotic distribution of $n^{1/2}(\hat{\theta}-\theta)$. Thus for large n the distribution of $\hat{\theta}$ around the unknown θ is the same as the bootstrap distribution of $\hat{\theta}^{(b)}$ around $\hat{\theta}$, which is revealed by a large number (B) of draws.

We now consider the application of the bootstrap to the specific case of the fixed effects estimates. Our discussion follows Simar (1992). Let the fixed effects estimates be $\hat{\beta}$ and $\hat{\alpha}_i$, from which we calculate $\hat{u}_i^*$ and $\hat{r}_i^*$ (i=1,...,N). Let the residuals be $\hat{v}_{it} = y_{it} - \hat{\alpha}_i - x'_{it}\hat{\beta}$ (i=1,...,N, t=1,...,T). The bootstrap samples will be drawn by resampling these residuals, because the v_{it} are the quantities analogous to the z's in the previous paragraph,

in the sense that they are assumed to be i.i.d., and the $\hat{v}_{it}$ are the observable versions of the v_{it}. (The sample size n above corresponds to NT.) So, for bootstrap iteration b(= 1,..,B) we calculate the bootstrap sample $\hat{v}_{it}^{(b)}$ and the pseudo data $y_{it}^{(b)} = \hat{\alpha}_i + x'_{it}\hat{\beta} + \hat{v}_{it}^{(b)}$. From these data we get the bootstrap estimates $\hat{\beta}^{(b)}$, $\hat{\alpha}_i^{(b)}$, $\hat{u}_i^{*(b)}$, and $\hat{r}_i^{*(b)}$, and the bootstrap distribution of these estimates is used to make inferences about the parameters.

We note that the estimates $\hat{u}_i$ and $\hat{r}_i$ depend on the quantity $\max_{j=1,..,N} \hat{\alpha}_j$. Since "max" is not a smooth function, it is not immediately apparent that this quantity is asymptotically normal, and if it were not the validity of the bootstrap would be in doubt. A rigorous proof of the validity of the bootstrap for this problem is given by Hall, Härdle and Simar (1995). They prove the equivalence of the following three statements: (i) $\max_{j=1,..,N} \hat{\alpha}_j$ is asymptotically normal. (ii) The bootstrap is valid as $T \to \infty$ with N fixed. (iii) There are no ties for $\max_{j=1,..,N} \alpha_j$; that is, there is a unique index i such that $\alpha_i = \max_{j=1,..,N} \alpha_j$. There are two important implications of this result. First, the bootstrap will not be reliable unless T is large. Second, this is especially true if there are near ties for $\max_{j=1,..,N} \alpha_j$, in other words, when there is substantial uncertainty about which firm is best.

We now turn to specific bootstrapping procedures, which differ in the way they draw inferences based on the bootstrap estimates. In each case, suppose that we are trying to construct a confidence interval for $u_i^* = \max_{j=1,..,N} \alpha_j - \alpha_i$. That is, for a given confidence level c, we seek lower and upper bounds L_i, U_i such that $P[L_i \le u_i^* \le U_i] = 1\text{-}c$. This statement should hold exactly for large T, and for small T it will be inaccurate to an unknown extent.

The simplest version of the bootstrap for the construction of confidence intervals is the *percentile bootstrap*. Here we simply take L_i and U_i to be the upper and lower c/2 fractiles of the bootstrap distribution of the $\hat{u}_i^{*(b)}$. More formally, let $\hat{F}$ be the bootstrap cumulative distribution function for $\hat{u}_i^*$, so that $\hat{F}(s) = P(\hat{u}_i^{*(b)} \leq s)$ = the fraction of the B bootstrap replications in which $\hat{u}_i^{*(b)} \leq s$. Then we take $L_i = \hat{F}^{-1}(c/2)$ and $U_i = \hat{F}^{-1}(1\text{-}c/2)$.

The percentile bootstrap intervals are accurate for large T but may be inaccurate for small to moderate T. This is a general statement, but in the present context there is a specific reason to be worried, which is the finite sample upward bias in $\max_{j=1,\ldots,N} \hat{\alpha}_j$ as an estimate of $\max_{j=1,\ldots,N} \alpha_j$. This will be reflected in improper centering of the intervals and therefore inaccurate coverage probabilities. Simulation evidence on the severity of this problem is given in the following chapter and also given by Hall, Härdle and Simar (1993). Several more sophisticated (or at least more complicated) versions of the bootstrap have been suggested to construct more accurate confidence intervals. Hall, Härdle and Simar (1993, 1995) suggested the *iterated bootstrap*, also called the double bootstrap, which consists of two stages. The first stage is the usual percentile bootstrap, which constructs, for any given c, a confidence interval that is intended to hold with probability 1-c. We will call these "nominal" 1-c confidence intervals. The second stage of the bootstrap is used to estimate the true coverage probability of the nominal 1-c confidence intervals, as a function of c. That is, if we define the function $\pi(c)$ = true coverage probability level of the nominal 1-c level confidence interval from the percentile bootstrap, then we attempt to evaluate the function $\pi(c)$. When we have done so, we find c^*, say, such that $\pi(c^*) = 1\text{-}c$, and then we use as our confidence interval the

nominal $1-c^*$ level interval from the first stage percentile bootstrap, which we "expect" to have a true coverage probability of 1-c.

The mechanics of the iterated bootstrap are uncomplicated but time-consuming. For each of the original (first stage) bootstrap iterations B, the second stage involves a set of B_2 draws from the bootstrap residuals, construction of pseudo data, and construction of percentile confidence intervals, which then either do or do not cover the bootstrap estimate $\hat{\theta}^{(b)}$. The coverage probability function $\pi(c)$ is the fraction of times coverage occurs. Generally we take $B_2 = B$, so that the total number of draws has increased from B to B^2 by going to the iterated bootstrap. Theoretically, the error in the percentile bootstrap is of order $n^{-1/2}$ while the error in the iterated bootstrap is of order n^{-1}. There is no clear connection between this statement and the question of how well finite sample bias is handled.

An objection to the iterated bootstrap is that it does not explicitly handle bias. If the nominal 90% confidence intervals only cover 75% of the time, it simply insists on a higher nominal confidence level, like 98%, so as to get 90% coverage. That is, it just makes the intervals wider, when bias might more reasonably be handled by recentering the intervals. A technique that does recenter the intervals is the *bias-adjusted bootstrap* of Efron (1982, 1985). As above, let θ be the parameter of interest, $\hat{\theta}$ the sample estimate and $\hat{\theta}^{(b)}$ the bootstrap estimate (for b=1,...,B), and let $\hat{F}$ be the bootstrap cdf. For n large enough that the bootstrap is accurate, we should expect $\hat{F}(\hat{\theta}) = 0.5$, and failure of this to occur is a suggestion of bias. Now define $z_0 = \Phi^{-1}[\hat{F}(\hat{\theta})]$ where Φ is the standard normal cdf, and where $\hat{F}(\hat{\theta}) = 0.5$ would imply $z_0 = 0$. Let $z_{c/2}$ be the usual normal critical value;

e.g. for c = 0.05, $z_{c/2} = z_{.025} = 1.96$. Then the bias-adjusted bootstrap confidence interval is $[L_i, U_i]$ with:

(2.8) $$L_i = \hat{F}^{-1}[\Phi(2z_0 - z_{c/2})], \; U_i = \hat{F}^{-1}[\Phi(2z_0 + z_{c/2})].$$

For example, suppose that there is upward bias, reflected by the fact that 60% of the bootstrap draws are larger than $\hat{\theta}$, so that $\hat{F}(\hat{\theta}) = 0.4$. Then $z_0 = -0.253$, and for c = 0.05 we have $\Phi(2z_0 - z_{c/2}) = \Phi(-2.466) = 0.0068$ and $\Phi(2z_0 + z_{c/2}) = 0.937$. Thus our confidence interval comes from the lower tail 0.0068 fractile and the upper tail 0.063 fractile, and we have compensated for upward bias by moving the interval left. This seems intuitively reasonable.

The assumption that justifies the bias-adjusted bootstrap is that, for some monotone increasing function g, $[g(\hat{\theta}) - g(\theta)]$ is distributed as $N(-z_0\sigma, \sigma^2)$ and $[g(\hat{\theta}^{(b)}) - g(\hat{\theta})]$ is also distributed as $N(-z_0\sigma, \sigma^2)$, for some z_0, σ^2. (The first distribution is from the probability law of the sample and the second is the bootstrap probability distribution induced by resampling from the given sample.) Thus we have normality, and also equal biases and variances, for some transformation of θ. The transformation function g need not be known. This is an advantage in implementation, but a disadvantage in trying to decide whether the assumption holds. It is not known whether the bias-adjusted bootstrap is valid for our specific problem, but it performs relatively well in the simulations reported in Chapter 3.

The final version of the bootstrap that we will consider is the *bias-adjusted and accelerated bootstrap* of Efron and Tibshirani (1993). This is intended to allow for the

possibility that the variance of $\hat{\theta}$ depends on θ, so that a bias-adjustment also requires a change in variance. This correction depends on some quantities defined in terms of the so-called jackknife values of $\hat{\theta}$. For i=1,...,n, let $\hat{\theta}_{(i)}$ be the value of the estimate based on all observations other than observation i; and let $\hat{\theta}_{(\bullet)} = n^{-1} \sum_{i=1}^{n} \hat{\theta}_{(i)}$ be the average of these values. Then the "acceleration" factor a is defined by:

$$(2.9) \qquad a = \sum_{i=1}^{n} (\hat{\theta}_{(\bullet)} - \hat{\theta}_{(i)})^3 / 6[\sum_{i=1}^{n} (\hat{\theta}_{(\bullet)} - \hat{\theta}_{(i)})^2]^{1.5}$$

With z_0 and $z_{c/2}$ defined as above, define

$$(2.10) \qquad b_1 = z_0 + (z_0 + z_{c/2})/[1 - a(z_0 + z_{c/2})] \,, \; b_2 = z_0 + (z_0 + z_{1-c/2})/[1 - a(z_0 + z_{1-c/2})]$$

Then the confidence interval is $[L_i, U_i]$, with $L_i = \hat{F}^{-1}[\Phi(b_1)]$ and $U_i = \hat{F}^{-1}[\Phi(b_2)]$. More discussion can be found in Efron and Tibshirani (1993, chapter 14).

3. Bayesian Procedures

In this section we discuss the Bayesian analysis of the stochastic frontier model. Bayesian analyses have been proposed and described in a series of papers by Koop, Osiewalski and Steel (hereafter KOS), especially KOS (1997) but also including Broeck, Koop, Osiewalski and Steel (1994), Koop, Steel and Osiewalski (1995), and Osiewalski and Steel (1998).

3.1 General Discussion

The basic Bayesian principles are straightforward. We have a set of observable data $Y = (y_1,...,y_n)$ and a vector θ (say of dimension K) of unobservable parameters. Let $p(\theta)$ be the prior density of θ and $p(Y \mid \theta)$ be the likelihood, where the prior is specified by the data analyst and the likelihood follows from the assumed model. Then Bayes Law says that:

$$(2.11) \qquad p(\theta|Y) \propto p(\theta)p(Y|\theta)$$

where $p(\theta|Y)$ is the posterior density of θ and "$\propto$" indicates proportionality. The traditional interpretation is that both the prior and the posterior reflect subjective probability distributions of θ, one (the prior) prior to the observation of Y and the other (the posterior) after the observation of Y. Bayes Law shows how the subjective probability distribution of θ is modified by the observation of Y. The concept of subjective probability is controversial but Bayes Law itself is not, since it is just the usual rule for conditional probability.

Inference on the parameters is performed using the posterior distribution. Since θ is usually multidimensional, one must face the often considerable problem of obtaining the marginal posterior distribution for a single given parameter such as θ_i (for some specific value $1 \leq i \leq K$). The marginal posterior density of θ_i is in principle defined by integrating the joint posterior density of θ with respect to all elements of θ other than θ_i, but this integral may not be analytically tractable. An alternative is to make Monte Carlo draws from the posterior distribution $p(\theta|Y)$ and to use these to reveal whatever features of the distribution of θ_i are interesting. Some numerical problems related to such Monte Carlo procedures will be discussed below.

Bayesian methods treat the parameters as random and condition on the data, which is more or less exactly the opposite of what classical methods do. However, in the present context of the stochastic frontier model with panel data, these distinctions can become a bit blurred. The parameters will typically include α, β, σ_v^2, $u_1,...,u_N$, and possibly some additional nuisance parameters. Existing classical treatments of this model have always treated α, β and σ_v^2 as fixed, but the inefficiency terms $u_1,...,u_N$ have often been treated as random and sometimes assigned a distribution. As discussed in section 2.1 above, inference on u_i is then performed using the distribution of u_i conditional on $(\varepsilon_{i1},....,\varepsilon_{iT})$, which is certainly similar to a posterior distribution. In fact this can be regarded as Bayesian inference in classical clothing. It differs from the Bayesian posterior in that it treats α, β and the nuisance parameters in the distributions of v_{it} and u_i as known, whereas the Bayesian posterior conditions only on the data. This difference is not likely to be substantial in practice because the parameters being taken as known are estimated based on NT observations (rather than just T observations for u_i) and should not contribute much variability to the Bayesian posterior.

3.2 The Bayesian Fixed Effects Model

In this section we discuss a model that KOS call the *standard individual effects* model (or SIE model). They regard it as one possible variant of the *Bayesian fixed effects* model, whereas we will just refer to it as the Bayesian fixed effects model, but this is only a semantic point.

This model postulates an "uninformative" prior for the basic parameters $\alpha_1,...,\alpha_N$, β, σ_v^2: $p(\alpha_1,...,\alpha_N, \beta, \sigma_v^2) \propto \sigma_v^{-2}$. (We do not regard this prior as

uninformative, but again this is just a semantic point.) Note that, in contrast to random effects models to be discussed later, we do not attempt to identify α and $u_1,...,u_N$ separately. Rather we simply measure relative inefficiency, by considering $u_i^* = \max_{j=1,..N} \alpha_j - \alpha_i$ and $r_i^* = \exp(-u_i^*)$ as functions of the firm-specific intercepts. This is similar in spirit to the classical fixed-effects treatment.

The likelihood $p(Y|\alpha_1,...,\alpha_N, \beta, \sigma_v^2)$ is the usual (classical) normal likelihood that would follow from treating the x_{it} as fixed and the v_{it} as i.i.d. normal. Specification of the prior and the likelihood defines the problem and implies the form of the posterior. The marginal posterior of $\alpha_1,...,\alpha_N, \beta$ can be calculated analytically to be (N+K)-variate student t with N(T-1)-K degrees of freedom. For any reasonable problem the number of degrees of freedom is large enough to treat the posterior as multivariate normal. The posterior mean of β is the classical fixed-effects estimate $\hat{\beta}$ as in section 2.2, and similarly the posterior mean of α_i is the fixed-effects estimate $\hat{\alpha}_i = \bar{y}_i - \bar{x}_i'\hat{\beta}$. It is also the case that the posterior variance matrix for β and $\alpha_1,...,\alpha_N$ is the same as the classical result for the variance matrix of $\hat{\beta}$ and $\hat{\alpha}_1,...,\hat{\alpha}_N$. For all of these reasons the name Bayesian fixed effects model seems appropriate.

The posterior distribution of the inefficiency estimate u_i^* or r_i^* is potentially complicated, but is easily revealed by Monte Carlo draws from the multivariate normal posterior distribution of $\alpha_1,...,\alpha_N$. For example, confidence intervals are easily constructed from the percentiles of these draws. These are the same confidence intervals that would be constructed by a classical econometrician via a simulation from the

estimated distribution of $\hat{\alpha}_1,\ldots,\hat{\alpha}_N$. We suspect that they will also often be similar to the confidence intervals constructed by bootstrapping the fixed effects estimates. They will differ only to the extent that the empirical distribution of the residuals $\hat{v}_{it}$ is not similar to the distribution of i.i.d. normals (which would reflect a failure of the assumed model).

An important point, stressed by KOS (1997), is that the fixed effects model favors low efficiency. An uninformative (flat) prior for $\alpha_1,\ldots,\alpha_N$ implies an uninformative (flat) prior for u_i^*, but an informative prior for $r_i^* = \exp(-u_i^*)$. More specifically, if u has constant density on $[0,\infty)$, then $r = \exp(-u)$ has density proportional to r^{-1} on $(0,1]$. This is an improper (prior) density that loosely speaking puts infinitely more weight on low values of r than high ones; for any constant c in (0,1), no matter how small, there is infinite weight on r in (0,c) but finite weight on c in (c,1]. One could argue about whether this reflects a problem with this specific prior, or with improper priors in general, but in any case it implies that we should expect the Bayesian fixed effects model to yield smaller posterior efficiencies than a model with more or less any (proper) informative prior. In a sense this fact is the Bayesian counterpart to the finite sample bias problem discussed in section 2.2. Whether treated in a classical or Bayesian way, the fixed effects model will tend to yield smaller efficiency values than models that assert a distribution for inefficiency.

3.3 Bayesian Random Effects Models

In this section we consider models that have an informative prior for u_i. This allows us to distinguish u_i from the overall intercept α, and so now we can estimate absolute inefficiency (u_i) instead of just relative inefficiency (u_i^*). Thus the parameters of

the problem are α, β, σ_v^2, $u_1,...,u_N$ and ϕ, where ϕ (which is present only in some models) represents parameters in the distribution of u. In all cases we take the likelihood $p(Y|\theta)$ to be the same normal likelihood as in the fixed effects case. In all cases we use the uninformative prior for α, β and σ_v^2: $p(\alpha, \beta, \sigma_v^2) \propto \sigma_v^{-2}$. When ϕ does not exist, $u_1,...,u_N$ is prior independent of α, β, σ_v^2. When ϕ does exist, ϕ and $u_1,...,u_N|\phi$ are prior independent of α, β, σ_v^2.

The models we consider all assert in one way or another that u follows an exponential distribution with mean λ, so that $p(u_i|\lambda) = \lambda^{-1}\exp(-u_i/\lambda)$. They differ in how λ is treated.

The first model we consider postulates an uninformative prior for r. More precisely, the r_i (i=1,...,N) are i.i.d. as uniform on (0,1]. This uninformative prior for r implies an informative prior for $u = -\ln r$; the u_i are i.i.d. with density proportional to $\exp(-u_i)$, so that u_i is exponential with $\lambda = 1$. For this model, because the value of λ is specified, there are no nuisance parameters (ϕ) in the distribution of u.

Our second model differs from the first only because it uses a different value of λ. The first model implied prior median efficiency of 0.5, which seems low for at least some applications. The second model chooses λ to imply prior median efficiency of 0.8. This is achieved by picking $\lambda = -\ln(0.8)/\ln(2) = 0.322$.

Our last two models differ from the first two in that λ is now treated as a parameter. That is, we specify a hierarchical prior in which conditional on λ the u_i are i.i.d. as exponential with mean parameter λ, and then we specify a prior for λ. Thus "ϕ" in the notation above now corresponds to λ, the nuisance parameter in the distribution of u_i.

Whereas the u_i are mutually prior independent conditional on λ, their dependence on a common value of λ implies that unconditionally they are not prior independent. KOS (1997, p. 86) refer to this as the "common efficiency distribution" or CED model.

For our third model, we specify an "uninformative" prior for λ: $p(\lambda) \propto \lambda^{-1}$. For our fourth model, we follow KOS (1997) and assume that λ^{-1} is exponential with mean $-1/\ln(r_{med})$, where r_{med} is the specified prior median efficiency. We take $r_{med} = 0.8$ as above (whereas they used 0.85). As noted by KOS, this hierarchical prior implies that the prior distribution of u_i is three-parameter inverted beta, but this model differs from the model (which they call the MIED model) in which the u_i are i.i.d. as three-parameter inverted beta, because in the present model the u_i are dependent due to the common value of λ.

For each of the above models, the specification of the prior and of the form of the likelihood implies the form of the posterior, $p(\alpha,\beta,\sigma_v^2,u_1,..,u_N,\phi|Y)$. In principle, inference on u_i would be conducted based on its marginal posterior, $p(u_i|Y)$, but the integrals needed to construct the marginal posterior analytically are intractable. Numerical integration techniques such as those used by Broeck et al. (1994) in the cross-sectional case are likely to be impractical in the present setting due to the dimensionality of the integral. We will follow Koop, Steel and Osiewalski (1995), KOS (1997) and Osiewalski and Steel (1998) in using Gibbs Sampling to make Monte Carlo draws from the joint posterior. These draws then reveal the posterior distribution of the parameters such as u_i; in particular, confidence intervals for u_i are easily constructed from the percentiles of the Monte Carlo draws. This is the same principle as was followed for the

Bayesian fixed effects model, except that there Gibbs Sampling was unnecessary because of the simple form of the joint posterior.

Gibbs sampling is a general procedure that makes draws from a joint distribution by making iterated sequential draws from the conditional distributions. It is useful in cases like the present one in which the conditional distributions are much simpler than the joint distribution, so that we know how to make draws from them. We split the set of parameters into three subsets: $(\alpha,\beta,\sigma_v^2)$, $(u_1,..,u_N)$ and ϕ. Starting from some arbitrary starting values, say $(\alpha,\beta,\sigma_v^2)^{(0)}$, $(u_1,..,u_N)^{(0)}$ and $\phi^{(0)}$, we generate random draws in sequence from the conditional distributions, and then we iterate. Thus, at step j, make the following draws:

$(\alpha,\beta,\sigma_v^2)^{(j)}$ from $p(\alpha,\beta,\sigma_v^2|Y,(u_1,..,u_N)^{(j-1)},\phi^{(j-1)})$

$(u_1,..,u_N)^{(j)}$ from $p(u_1,..,u_N|Y,(\alpha,\beta,\sigma_v^2)^{(j-1)},\phi^{(j-1)})$

$\phi^{(j)}$ from $p(\phi|Y,(\alpha,\beta,\sigma_v^2)^{(j-1)},(u_1,..,u_N)^{(j-1)})$.

For large enough s that convergence has occurred, $(\alpha,\beta,\sigma_v^2)^{(s)}$, $(u_1,..,u_N)^{(s)}$ and $\phi^{(s)}$ can be treated as draws from the joint posterior. The reader is referred to Dorfman (1997) for more discussion of Gibbs Sampling, and to KOS (1997, Appendix) for the forms of the conditional distributions needed in the present case.

4. Empirical Results

We now proceed to apply the classical and Bayesian procedures described above to three previously-analyzed data sets. These data sets were chosen to have rather different characteristics. The first data set consists of N=171 Indonesian rice farms

observed for T=6 growing seasons. We have $\sigma_v^2 = 0.108$ and $\sigma_u^2 = 0.007$ (these values being the exponential MLE's). Inference on inefficiencies will be very imprecise because T is small and because σ_v^2 is large relative to σ_u^2. The second data set consists of N=10 Texas utilities observed for T=18 years, with $\sigma_v^2 = 0.003$ and $\sigma_u^2 = 0.020$. For this data set we can estimate inefficiencies much more precisely because T is larger and σ_v^2 is smaller relative to σ_u^2. The third data set consists of N=25 Egyptian tileries observed for a maximum of T=22 production periods, with $\sigma_v^2 = 0.113$ and $\sigma_u^2 = 0.057$. This is a case that is intermediate between the other two. We will see that the precision of estimation of the efficiency levels will indeed differ strikingly across these data sets, and that choice of technique will matter more where precision is low.

4.1 Indonesian Rice Farms

These data are due to Erwidodo (1990) and have been analyzed subsequently by Lee (1991), Lee and Schmidt (1993), Schmidt and Horrace (1996, 1999) and others. There are N=171 rice farms and T=6 six-month growing seasons. Output is rice in kilograms and inputs are land, labor, seed and two types of fertilizer. The functional form is Cobb-Douglas with some dummy variables added for region, seasonality, and some types of farming practices. For a complete discussion of the data see Erwidodo (1990).

Table 1 gives point estimates for the regression parameters for the fixed effects model (within estimates); for the classical MLE's based on the half-normal and exponential distributions for u; and for our four Bayesian models. For the Bayesian models the point estimates are the posterior means and the standard errors are the square

roots of the posterior variances. Here and subsequently the heading "uninformative" refers to our third Bayesian model, with an uninformative prior for the exponential parameter; "hierarchical" refers to our fourth model, with an exponential prior for the inverse of the exponential parameter; and "λ=.322" and "λ=1" refer to our second and first models, with exponential priors with specified values of λ, so as to imply median prior efficiencies of 0.8 and 0.5 respectively. The results in Table 1 are quite similar across techniques. If we were primarily interested in the regression parameters, as opposed to the firm-specific efficiency levels, it would really not make much difference which technique we picked.

Table 2 gives point estimates and 90% confidence intervals for the relative efficiency measures r_i^*, for the classical and Bayesian fixed effects models. For the classical fixed effects model, we give the usual point estimate based on the within estimates; the MCB and MargCB confidence intervals; and confidence intervals based on three versions of the bootstrap. We used the percentile method and the iterated (two stage) percentile method, and also the bias-adjusted and accelerated bootstrap (labelled BC_a). The bias-adjusted bootstrap (without acceleration) gave very similar results to the bias-adjusted and accelerated bootstrap so we do not report them separately. We used B=1000 bootstrap replications, except that for the iterated bootstrap we used only $B=B_2=200$ replications, to shorten the computational time. (For the other two data sets, which are smaller, we used $B=B_2=1000$ replications for the iterated bootstrap.) For the Bayesian fixed-effects model, we report the posterior mean and the 90% confidence intervals based on the appropriate percentiles of the simulated posterior distribution.

There are 171 firms and so we report results only for a few of them. We report results for the three firms (164, 118 and 163) that are most efficient; for the firms at the 75th percentile (31), 50th percentile (15) and 25th percentile (16) of the efficiency distribution; and for the two worst firms (117, 45). All of these rankings are according to the classical fixed effects estimates.

In terms of the point estimates of efficiency levels, the classical and Bayesian fixed effects estimates are relatively similar. The Bayesian estimates are a little lower, especially for the most efficient firms. The efficiency estimates overall are rather low, with median efficiency only a little over 0.5. This is as expected.

We next discuss the confidence intervals for efficiency levels. The MCB and MargCB intervals are quite wide, especially for the less efficient firms. In fact, they really are too wide to be of much use. The bootstrap and Bayesian intervals are narrower, but still disappointingly wide. The percentile method of the bootstrap and the Bayesian fixed effects model give intervals that are quite similar. As noted above, this is as expected. With only T=6 observations per firm, the accuracy of the percentile bootstrap is suspect. No such statement applies to the Bayesian method, but with only T=6 observations per firm, the prior is certainly not dominated by the data, and since the prior is arguably unreasonable, so are the posterior results. As expected, the iterated bootstrap intervals are wider than those from the percentile method, and the bias-adjusted and accelerated bootstrap confidence intervals are shifted to the right (in the direction of higher efficiency levels). Simulations reported in Chapter 3 suggest that the bias-adjustment is helpful, but with so few observations per firm all of the bootstrap methods are probably unreliable.

Table 3 gives the estimated efficiencies and the associated 90% confidence intervals for our random effects models, including the classical MLE's based on the half normal and exponential distributions and our four Bayesian models with informative priors. We will not discuss the half normal results other than to note that they are not too different from the exponential. All of the confidence intervals are disappointingly wide, as they were for the fixed effects models of Table 2. However, comparing Tables 2 and 3, it is apparent that the efficiency levels from the random effects models of Table 3 are considerably higher than those from the fixed effects models. As noted above, from the classical point of view this is a reflection of the bias in the fixed effects efficiency estimates, while from a Bayesian point of view it reflects the influence of the underlying prior which is very heavily weighted toward low efficiency levels.

The exponential MLE and the Bayesian model with an uninformative prior for the exponential parameter give strikingly similar results. It is not surprising that the results are similar but it is perhaps unexpected that they are so similar. The Bayesian model with $\lambda=1$ has a lower prior median efficiency (equal to 0.5) than the hierarchical model or the model with $\lambda=0.322$ (both of which have prior median efficiency of 0.8) and correspondingly has lower posterior efficiencies. The choice of prior matters a fair amount, which is a reflection of the small amount of data ($T=6$) per firm.

4.2 Texas Utilities

We next consider the Texas utility data of Kumbhakar (1996), which was also analyzed by Horrace and Schmidt (1996, 1999). There are $N=10$ privately owned Texas electric utilities observed for $T=18$ years. Kumbhakar estimated a cost function, whereas

we will estimate a Cobb-Douglas production function. Output is electric power generated and the inputs are measures of labor, capital and fuel. For more details on the data see Kumbhakar (1996).

Table 4 gives the regression parameter estimates. We will not comment on these except to note that the variance of the one-sided error is large relative to the variance of noise (e.g. 0.020 vs. 0.003 for the exponential MLE). This is the opposite of the case for the Indonesian rice farms. For this reason, and because T is larger here (18 vs. 6), we expect more precise estimates of efficiency levels and less sensitivity of the results to the choice of method for this data set than for the previous one.

Table 5 gives the estimated efficiencies and 90% confidence intervals for the fixed effects models, while Table 6 gives the same results for the random effects models. The format is the same as for Tables 2 and 3, except that we can display results for all N=10 firms. We can see the same patterns here as we did for the previous data set (though less distinctly since differences across techniques are smaller). The MCB and MargCB intervals are wider than the other intervals. The fixed effects models give lower estimates of efficiency levels than the random effects models. Comparable classical and Bayesian models given comparable results: the classical fixed effects results with the percentile bootstrap are quite similar to the Bayesian fixed effects results, and the results from the classical exponential MLE are quite similar to those from the Bayesian model with an uninformative prior for the exponential parameter. Bayesian models with higher prior efficiency levels yield higher posterior efficiency levels.

However, we repeat that the main result of interest is the general comparison of the results for this data set with those from the previous data set. For this data set we can

estimate efficiency levels precisely enough to make reasonable statements about them, and the choice of technique is not critically important. The most important aspect of the choice of technique is the choice of a fixed versus random effects model. The fixed effects models give lower efficiencies, and are suspect for the reasons discussed in sections 3.2 and 4.2 above. The gain from being willing to assert a distribution for inefficiency is large, and at least for this data set there is reasonable robustness to the choice of distribution.

4.3 Egyptian Tileries

The last data set we consider is for Egyptian tileries. It was collected by Seale (1990) and has been analyzed by Horrace and Schmidt (1996, 1999). There are N=25 small-scale Egyptian manufacturers of ceramic floor tiles, with observations for a maximum of T=22 three-week production periods. There are some missing data points (production did not occur in some periods) and so this is an unbalanced panel. Output is square meters of tile, while inputs are labor and machine hours.

The results are given in Tables 7, 8 and 9. As with the Indonesian rice farms, we present results only for a subset of the firms, choosing firms at the same percentiles of the efficiency distribution as in section 4.1. We will not discuss these results in detail, but the same comparisons across techniques that held for the other two data sets hold here as well. Comparing results across data sets, the results here are more precise, and less dependent on choice of technique, than for the Indonesian rice farms; they are less precise, and more dependent on choice of technique, than for the Texas utilities. This is predictable because, both in terms of the sizes of N and T, and also in terms of the

relative variances of noise and inefficiency, this data set has characteristics that are intermediate between those of the previous two data sets.

5. Concluding Remarks

In this chapter we considered a large number of classical and Bayesian procedures to estimate technical efficiency levels of firms and to construct confidence intervals for these efficiency levels. We then applied these methods to three data sets with different characteristics that determined the difficulty of the estimation problem. Comparing results across data sets and across methods within a data set leads to some clear and important conclusions.

First, the estimation problem is easier when T is large and when the variance of noise is small relative to the variance of inefficiency, and harder when T is small and when the variance of noise is large relative to the variance of inefficiency. In easier problems we can estimate efficiency levels more precisely than in harder problems, and there is less sensitivity of the results to the choice of technique.

Second, we do not find much difference between classical and Bayesian methods if we match methods that depend on comparable assumptions. For example, the Bayesian fixed effects model gives similar results to those obtained by the percentile bootstrap applied to the fixed effects (within) estimates. As another example, the classical MLE based on the exponential distribution gives similar results to the Bayesian model in which the prior distribution for inefficiency is exponential, and there is an uninformative prior for the exponential parameter. Furthermore, the two approaches face a similar problem, in that the results may be "unreliable" when T is small. In the classical framework,

"unreliable" means that asymptotically valid inference may not be valid in small samples, while in the Bayesian framework "unreliable" means that the prior will not be dominated by the data and so there is a lack of robustness to the choice of the prior. We do *not* mean to allege that these are the same problem, but simply that in either approach small T causes problems.

Third, the multiple comparisons with the best (MCB) and marginal comparisons with the best (MargCB) intervals are wider than any of the others we consider. This is not a desirable feature in a confidence intervals, but on the other hand these intervals are valid for small T, and a conservative, valid interval at least can provide the correct message that in some cases we don't know very much.

Finally, the main difference in results is between fixed effects and random effects models. Fixed effects models (either classical or Bayesian) yield much lower efficiency levels than random effects models, and there are good reasons to be skeptical of the fixed effects results. From a classical point of view, the fixed effects estimates of efficiency levels are biased downward; from the Bayesian point of view, the fixed effects model embodies a prior that is unreasonably heavily weighted toward low efficiency levels. Random effects models require a distributional assumption for inefficiency, which may be unattractive. However, making such an assumption yields large dividends in terms of precision of estimation *and* in terms of more reasonable average levels of efficiency.

Table 1

Estimates of Parameters : Indonesian Rice Farm

Variables	FE Model	RE model					
		MLE		Bayesian			
		Half-normal	Exponential	Uninformative	Hierarchical	λ=.322	λ=1
Constant		5.199 (.194)	5.181 (.193)	5.187 (.193)	5.187 (.195)	5.319 (.206)	5.419 (.220)
Seed	.121 (.030)	.134 (.027)	.135 (.027)	.135 (.026)	.135 (.027)	.129 (.028)	.125 (.029)
Urea	.092 (.021)	.113 (.018)	.113 (.018)	.113 (.018)	.113 (.018)	.102 (.019)	.096 (.020)
TSP	.089 (.013)	.076 (.012)	.076 (.011)	.076 (.012)	.077 (.011)	.084 (.012)	.088 (.012)
Labor	.243 (.032)	.219 (.029)	.217 (.029)	.217 (.029)	.217 (.029)	.224 (.031)	.234 (.031)
Land	.452 (.035)	.481 (.031)	.483 (.031)	.483 (.030)	.483 (.031)	.477 (.033)	.466 (.035)
DP	.034 (.032)	.009 (.029)	.008 (.028)	.008 (.029)	.008 (.029)	.013 (.029)	.020 (.031)
DV1	.179 (.041)	.176 (.038)	.176 (.038)	.176 (.038)	.176 (.039)	.177 (.038)	.177 (.040)
DV2	.175 (.057)	.140 (.052)	.136 (.052)	.138 (.052)	.138 (.052)	.149 (.054)	.161 (.055)
DSS	.053 (.022)	.049 (.021)	.049 (.021)	.050 (.022)	.050 (.022)	.050 (.021)	.052 (.021)
DR1		- .058 (.049)	-.059 (.048)	-.060 (.049)	-.060 (.050)	-.093 (.072)	-.125 (.109)
DR2		- .047 (.058)	-.045 (.057)	-.046 (.057)	-.046 (.059)	-.075 (.076)	-.112 (.105)
DR3		- .078 (.062)	-.077 (.060)	-.078 (.062)	-.078 (.062)	-.126 (.080)	-.176 (.110)
DR4		.016 (.058)	.021 (.056)	.018 (.057)	.020 (.059)	-.004 (.082)	-.025 (.117)
DR5		.082 (.060)	.089 (.059)	.088 (.059)	.090 (.061)	.092 (.079)	.085 (.110)
σ_v^2	.108	.108 (.005)	.108 (.005)	.110 (.005)	.110 (.005)	.106 (.005)	.106 (.005)
σ_u^2		.007 (.003)	.007	.007	.006		
λ			.081 (.019)	.084 (.019)	.089 (.018)		

Table 2

Estimates of Efficiencies (Fixed Effect Model) : Indonesian Rice Farm

Firm No.	FE estim-ates	MCB				Bootstrap						Bayesian SIE		
		Simultaneous		Marginal		Percentile method		BC_a		Iterated bootstrap				
		LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	Mean	LB	UB
164	1	.583	1	.737	1	.748	1	.781	1	.712	1	.912	.717	1
118	.932	.508	1	.642	1	.681	1	.772	1	.643	1	.854	.657	1
163	.930	.509	1	.643	1	.674	1	.773	1	.642	1	.850	.653	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
31	.616	.328	1	.421	1	.441	.724	.529	.878	.420	.756	.565	.423	.731
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
15	.554	.300	1	.379	1	.398	.646	.471	.746	.379	.673	.509	.383	.656
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
16	.498	.266	1	.340	1	.358	.589	.423	.679	.342	.614	.458	.344	.590
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
117	.379	.203	1	.259	.974	.272	.445	.321	.517	.259	.465	.348	.258	.450
45	.365	.197	1	.250	.940	.262	.427	.310	.496	.249	.445	.336	.253	.432

Table 3

Estimates of Efficiencies (Random Effect Model) : Indonesian Rice Farm

Firm No.	MLE						Bayesian											
	Half-normal			Exponential			Uninformative			Hierarchical			λ=.322			λ=1		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
164	.964	.903	.998	.973	.924	.999	.974	.923	.999	.973	.921	.999	.954	.871	.997	.931	.820	.995
118	.964	.902	.998	.973	.923	.999	.973	.921	.999	.972	.918	.998	.958	.881	.997	.936	.827	.996
163	.959	.889	.998	.970	.924	.999	.969	.912	.998	.968	.908	.998	.946	.851	.997	.912	.785	.993
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
31	.924	.823	.994	.949	.863	.997	.948	.854	.997	.946	.853	.997	.891	.750	.991	.828	.657	.975
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
15	.923	.792	.990	.935	.834	.996	.935	.823	.996	.931	.819	.995	.820	.655	.972	.696	.529	.884
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
16	.845	.725	.969	.886	.751	.990	.885	.729	.991	.879	.722	.988	.719	.560	.897	.611	.463	.780
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
117	.773	.658	.907	.792	.643	.956	.796	.623	.965	.785	.610	.958	.591	.463	.741	.496	.379	.633
45	.774	.659	.908	.795	.645	.957	.798	.621	.968	.788	.621	.958	.600	.473	.747	.517	.403	.652

Table 4

Estimates of Parameters : Texas Utilities

Variables	FE Model	RE model					
		MLE		Bayesian			
		Half-normal	Exponential	Uninformative	Hierarchical	λ=.322	λ=1
Constant		-5.053	-5.064	-5.045	-5.048	-5.049	-5.060
		(.245)	(.245)	(.249)	(.254)	(.252)	(.254)
Labor	-.129	-.078	-.076	-.080	-.082	-.084	-.087
	(.046)	(.031)	(.029)	(.031)	(.032)	(.031)	(.032)
Capital	.628	.586	.585	.588	.590	.592	.600
	(.053)	(.049)	(.048)	(.049)	(.049)	(.049)	(.050)
Fuel	.565	.584	.584	.583	.582	.581	.579
	(.039)	(.038)	(.038)	(.038)	(.038)	(.039)	(.039)
σ_v^2	.003	.003	.003	.003	.003	.003	.003
		(.0003)	(.0003)	(.0003)	(.0003)	(.0003)	(.0003)
σ_u^2		.010	.020	.029	.030		
		(.005)					
λ			.143	.171	.173		
			(.048)	(.061)	(.061)		

Table 5

Estimates of Efficiencies (Fixed Effect Model) : Texas Utilities

Firm No.	FE Estim-ates	MCB				Bootstrap						Bayesian SIE		
		Simultaneous		Marginal		Percentile method		BC_a		Iterated bootstrap				
		LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	Mean	LB	UB
5	1	.945	1	.960	1	.983	1	.983	1	.977	1	.997	.981	1
3	.916	.797	1	.806	1	.828	1	.832	1	.820	1	.914	.820	1
10	.861	.765	.984	.773	.979	.790	.926	.792	.927	.786	.928	.859	.784	.926
1	.835	.769	.984	.775	.976	.786	.876	.789	.878	.782	.878	.832	.783	.878
8	.820	.759	.973	.764	.964	.775	.858	.777	.860	.772	.860	.817	.771	.859
9	.806	.755	.972	.759	.961	.767	.840	.770	.843	.765	.842	.804	.766	.841
2	.801	.735	.942	.740	.934	.752	.843	.754	.845	.749	.845	.798	.748	.844
7	.786	.716	.920	.722	.913	.735	.831	.736	.832	.730	.834	.784	.731	.832
6	.786	.715	.917	.720	.910	.732	.832	.734	.833	.730	.834	.784	.729	.833
4	.762	.707	.910	.711	.901	.721	.798	.725	.802	.719	.800	.760	.719	.799

Table 6

Estimates of Efficiencies (Random Effect Model) : Texas Utilities

Firm No.	MLE						Bayesian											
	Half-normal			Exponential			Uninformative			Hierarchical			λ=.322			λ=1		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
5	.987	.971	.999	.989	.974	.999	.982	.950	.999	.981	.947	.999	.979	.943	.999	.977	.940	.998
3	.978	.959	.996	.983	.964	.998	.969	.923	.998	.966	.921	.997	.964	.919	.996	.958	.913	.994
10	.908	.889	.927	.913	.894	.932	.902	.862	.937	.899	.859	.936	.896	.859	.931	.891	.854	.927
1	.864	.846	.882	.867	.849	.886	.858	.828	.887	.856	.825	.886	.853	.825	.880	.848	.820	.876
8	.846	.828	.864	.849	.832	.867	.841	.811	.870	.838	.808	.868	.835	.807	.862	.831	.803	.859
9	.826	.809	.843	.828	.811	.846	.821	.790	.851	.818	.788	.849	.815	.787	.844	.812	.783	.840
2	.831	.814	.848	.834	.817	.852	.826	.795	.857	.823	.792	.854	.820	.790	.848	.816	.787	.844
7	.817	.800	.834	.820	.803	.837	.811	.779	.842	.810	.776	.842	.806	.776	.836	.802	.772	.833
6	.820	.803	.837	.823	.806	.841	.814	.782	.845	.812	.781	.843	.809	.779	.838	.804	.776	.833
4	.786	.770	.801	.789	.772	.805	.781	.751	.810	.779	.749	.808	.776	.748	.803	.771	.744	.799

Table 7

Estimates of Parameters : Egyptian Tileries

Variables	FE Model	RE model					
		MLE		Bayesian			
		Half-normal	Exponential	Uninformative	Hierarchical	λ=.322	λ=1
Constant		.856	.860	.871	.866	.916	1.031
		(.253)	(.246)	(.248)	(.250)	(.250)	(.276)
Labor	1.007	1.030	1.025	1.024	1.025	1.020	1.012
	(.040)	(.038)	(.038)	(.038)	(.038)	(.038)	(.039)
Machines	.042	.046	.046	.046	.046	.046	.045
	(.045)	(.033)	(.031)	(.031)	(.032)	(.033)	(.038)
σ_v^2	.115	.113	.113	.114	.114	.114	.115
		(.008)	(.007)	(.008)	(.008)	(.008)	(.008)
σ_u^2		.040	.057	.067	.065		
		(.014)					
λ			.239	.257	.254		
			(.056)	(.062)	(.060)		

Table 8

Estimates of Efficiencies (Fixed Effect Model) : Egyptian Tileries

Firm No.	FE Estim-ates	MCB				Bootstrap						Bayesian SIE		
		Simultaneous		Marginal		Percentile method		BC_a		Iterated bootstrap				
		LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	Mean	LB	UB
14	1	.739	1	.805	1	.782	1	.946	1	.754	1	.907	.783	1
24	.994	.724	1	.790	1	.775	1	.943	1	.752	1	.901	.775	1
25	.989	.724	1	.789	1	.771	1	.939	1	.745	1	.896	.769	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
18	.953	.670	1	.737	1	.734	1	.893	1	.708	1	.863	.732	1
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
4	.895	.648	1	.708	1	.692	.940	.850	1	.672	.967	.812	.688	.945
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
6	.645	.478	1	.518	.956	.505	.672	.619	.781	.491	.688	.585	.503	.671
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
7	.555	.416	.885	.451	.816	.436	.576	.536	.670	.423	.594	.502	.434	.577
8	.493	.358	.815	.392	.744	.380	.520	.469	.616	.368	.537	.447	.379	.521

Table 9

Estimates of Efficiencies (Random Effect Model) : Egyptian Tileries

Firm No.	MLE						Bayesian											
	Half-normal			Exponential			Uninformative			Hierarchical			λ=.322			λ=1		
	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB	Mean	LB	UB
14	.943	.865	.995	.948	.874	.996	.944	.864	.996	.945	.862	.996	.936	.847	.994	.902	.787	.988
24	.950	.877	.996	.951	.879	.996	.947	.867	.996	.947	.869	.996	.939	.851	.995	.902	.789	.988
25	.942	.863	.995	.946	.870	.996	.944	.863	.996	.943	.859	.996	.935	.847	.995	.896	.782	.987
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
18	.929	.844	.993	.934	.852	.994	.928	.829	.994	.930	.834	.994	.918	.816	.991	.874	.752	.981
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
4	.885	.787	.978	.896	.799	.984	.891	.782	.986	.893	.782	.986	.876	.764	.977	.825	.704	.950
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
6	.675	.600	.755	.662	.588	.745	.660	.573	.752	.662	.577	.757	.644	.561	.735	.599	.515	.683
:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:	:
7	.582	.517	.651	.567	.504	.638	.565	.492	.644	.568	.495	.646	.552	.484	.625	.514	.443	.591
8	.540	.477	.607	.513	.454	.581	.511	.438	.592	.513	.439	.595	.497	.427	.574	.460	.388	.534

Chapter 3

SIMULATION EVIDENCE ON THE ACCURACY OF INFERENCES IN STOCHASTIC FRONTIER MODELS

1. Introduction

In the previous chapter, we considered a large number of different models, and we discussed the relationship between the assumptions underlying the various models and their empirical results. We found that making a distributional assumption for inefficiency led to much tighter confidence intervals. However, it is important to know whether these confidence intervals are accurate. For example, it is clear empirically that MCB and MargCB confidence intervals are considerably wider than intervals based on bootstrapping the fixed effects estimates or based on a distributional assumption, but it is not known how conservative MCB and MargCB are, whether the bootstrapping or distribution-based estimates are reliable. In this chapter, we conduct Monte Carlo simulations to investigate the accuracy of the confidence intervals based on the methods of the previous chapter

Monte-Carlo analysis of the stochastic frontier model can be found in Olson, Schmidt and Waldman (1980), Coelli (1995), and Gong and Sickles (1989,1992). The first two papers mainly focus on the comparison between corrected OLS and MLE in the estimation of the regression parameters in a stochastic frontier model with cross-section data. They find that the performance of each method in estimating the constant term and the variance parameters depends on the relative size of the variance of inefficiency and

the sample size. Coelli extended the experiment to investigate the finite sample properties of several tests of the existence of inefficiency.

Gong and Sickles conducted a Monte Carlo analysis to examine the relative strengths of the stochastic frontier model and data envelopment analysis using panel data. They considered three alternative techniques in estimating the stochastic frontier model; MLE, corrected GLS, and the fixed effect estimator. The main focus was on investigating the robustness of the efficiency estimates over different functional forms and different underlying assumptions.

Our Monte-Carlo experiment is designed to investigate the relative performances of different methods in performing inference on efficiencies in the stochastic frontier model using panel data. The overall design of the experiment is similar to that of Olson et al. and Coelli. However, our study is differentiated from those two papers in that our study considers a panel data model and more importantly our main interest is in the efficiency estimates instead of the estimates of regression parameters. Our study also differs from Gong and Sickles's paper in two aspects. First, they evaluated the performances of the different estimation methods based on the correlation coefficients between true efficiencies and their estimates, and between true ranks and their estimates. These criteria are useful only when our interest is in the comparison of firms in a certain group. We focus on the evaluation of the each method in terms of the bias and accuracy of interval estimates (confidence intervals). This allows us to make a useful statement on the performance of each method in describing the efficiency level of a specific firm. The other difference lies in the estimation techniques of main interest. Even though Gong and Sickle considered three estimation techniques, inefficiency is estimated in the same way

in all three cases. A distribution of inefficiency was assumed to estimate frontier parameters in MLE and to correct the constant term in corrected GLS, but the information on the distribution of inefficiency was not used at all to estimate the firm specific inefficiencies. The firm specific inefficiency was just defined relative to the sample best from the mean over time of the estimated residuals, after the frontier parameters were estimated by the different techniques. Then, the difference of three techniques in estimating the inefficiency only comes from their relative performance in estimating the frontier parameters. Our interest here is in the comparison of techniques that estimate inefficiencies in fundamentally different ways. In particular, we want to compare estimation of inefficiencies using a distributional assumption to estimation of inefficiencies in terms of the comparison of a given firm to the sample best.

Section 2 briefly introduces the model on which our study is based. Section 3 describes the design of the experiments. The experimental results and their implications are discussed in section 4.

2. The Model

As in the previous chapter, we consider the basic panel data stochastic frontier model as follows.

$$(3.1) \qquad y_{it} = \alpha + x_{it}'\beta + \varepsilon_{it} \qquad \text{where } \varepsilon_{it} = v_{it} - u_i \quad \text{and} \quad u_i \geq 0$$

We want to focus our attention on two estimation methods, MLE and the fixed effect method (FE method hereafter). For ML estimation, the first thing to do is to assume a specific distribution of u_i in which case the u_i can be estimated in the absolute sense. The

most frequently used specifications for u_i, in both the theoretical and empirical literatures, are the half-normal distribution and the exponential distribution. The truncated normal distribution has also been frequently used as a generalization of the half normal distribution. The gamma distribution was suggested as a generalization of the exponential distribution, but its empirical application has been restricted by the complexity in its estimation procedure. Here, we choose a half- normal distribution due to its popularity as well as the simplicity of the estimation procedure. We have no evidence for other distributions, but there is no reason to think that our results depend critically on special features of the half normal distribution.

The point estimates and confidence intervals for inefficiencies can be calculated from the distribution of u_i conditional on ε_i following Battese and Coelli. If the distribution of u_i is $N(\mu, \sigma_u^2)$ truncated from the left at zero, the distribution of u_i, conditional on ε_i, is $N(\delta, \sigma_*^2)$ truncated from the left at zero, where δ and σ_*^2 are defined as follows.

$$(3.2) \qquad \delta = \frac{\mu\sigma_v^2 - T\sigma_u^2\bar{\varepsilon}}{\sigma_v^2 + T\sigma_u^2}, \quad \sigma_*^2 = \frac{\sigma_v^2\sigma_u^2}{\sigma_v^2 + T\sigma_u^2}$$

The half normal distribution corresponds to the case of $\mu=0$. To summarize, the estimation of u_i proceeds in two steps. First, we estimate the frontier parameters by MLE. Next, the point estimate and confidence intervals for u_i are calculated from $f(u_i|\varepsilon_i)$ evaluated at $\hat{\varepsilon}_i$. The point estimate is the mean of the conditional distribution, $E(u_i|\varepsilon_i)$, and the confidence intervals are based on the percentiles of the conditional distribution. We call this the MLE & BC method.

In the fixed effect model, technical inefficiency is regarded as a nonnegative fixed effect. The basic stochastic frontier model (3.1) is the standard panel data model with time-invariant individual effects.

(3.3) $$y_{it} = \alpha_i + x_{it}'\beta + v_{it} \qquad \text{where } u_i = \alpha - \alpha_i$$

α_i can be estimated without a distributional assumption as the mean residual in the i^{th} group, i.e., $\hat{\alpha}_i = \bar{y}_i - \bar{x}_i'\hat{\beta}$ where $\hat{\beta}$ is within estimator of β. $\hat{\alpha}_i$ consistently estimates α_i as $T \to \infty$. Inefficiencies can be estimated relative to the sample best, i.e., $\hat{u}_i^* = \max_{j=1,\ldots,N} \hat{\alpha}_j - \hat{\alpha}_i$. When we define technical inefficiencies in this way, the sample best firm is treated as 100% efficient. The justification for this comes from the fact that the sample maximum approaches the population maximum as $N \to \infty$, so that for large N inefficiencies can be measured relative to an absolute standard. For the confidence statements in the FE model, two methods are employed. One is the marginal MCB (MargCB) technique and the other is bootstrapping. The details of marginal MCB and bootstrapping can be found in Chapter 2 and Chapter 4. These two methods will be called FE & MargCB method, and FE & Bootstrap method.

3. Design of the Experiments

We consider the model with no regressors so that we can concentrate our interest on the estimation of efficiencies without having to be concerned about the nature of the regressors. In practical cases the regression parameters β are likely to be estimated so much more efficiently than the other parameters that treating them as known is not likely

to make much difference. Then, the parameter space reduces to $(\alpha, \sigma_v^2, \sigma_u^2$, N, T). Without loss of generality, we can fix the constant term α to any number, since a change in the constant term only shifts the estimated constant term by the same amount, without any effect on the bias and variance of any of the estimates. For simplicity, we fix the constant term equal to one.

We need two parameters to characterize the variance structure of model. It is natural to think in terms of σ_v^2 and σ_u^2. Alternatively, recognizing that σ_u^2 is the variance of the untruncated normal from which u is derived, not the variance of u, we can think instead in terms of σ_v^2 and var(u), where $\text{var(u)}=(\frac{\pi-2}{\pi})\sigma_u^2$. However, we obtain more readily interpretable results if we think instead in terms of the size of total variance and the relative allocation of total variance between v and u. The total variance is defined as $\sigma_\varepsilon^2=\sigma_v^2+\text{var(u)}$. Olson et al. used $\lambda=\frac{\sigma_u}{\sigma_v}$ to represent the relative variance structure, so that their parameterization was in terms of σ_ε^2 and λ. Coelli used σ_ε^2 and either $\gamma = \frac{\sigma_u^2}{\sigma_v^2+\sigma_u^2}$ or $\gamma^* = \frac{\text{var(u)}}{\sigma_v^2+\text{var(u)}}$. The choice among λ, γ, and γ^* is a matter of convenience. We decided to use γ^* due to its ease of interpretation, so that we will use the parameters σ_ε^2 and γ^*. The reason this is a convenient parameterization (compared to the "obvious" choice of σ_v^2 and σ_u^2) is that, following Olsen et al., one can show that comparisons among the various estimators are not affected by the value of total variance σ_ε^2. The effect of multiplying σ_ε^2 by a factor of "k" holding γ^* constant is as follows.

a) constant term: bias changes by a factor of $\sqrt{k}$ and variance changes by a factor of k

b) σ_v^2 and σ_u^2 : bias changes by a factor of k and variance changes by a factor of k^2

c) γ^* (or λ or γ): bias and variance are unaffected.

We set σ_ε^2 at 0.25, arbitrarily. Thus the only parameters left to consider in the experiment are N, T, and γ^*. We consider three values of γ^* to include a case in which the variance of v dominates, a case in which the variance of u dominates, and an intermediate case. We take $\gamma^* = 0.1$, 0.5, and 0.9 to represent the above three cases. With $\sigma_\varepsilon^2 = 0.25$, σ_v^2, σ_u^2 and var(u) are determined as follows for each value of γ^*.

a) $\gamma^* = 0.1$: $\sigma_\varepsilon^2 = 0.25$, $\sigma_v^2 = 0.225$, var(u) $= 0.025$, $\sigma_u^2 = 0.069$

b) $\gamma^* = 0.5$: $\sigma_\varepsilon^2 = 0.25$, $\sigma_v^2 = 0.125$, var(u) $= 0.125$, $\sigma_u^2 = 0.344$

c) $\gamma^* = 0.9$: $\sigma_\varepsilon^2 = 0.25$ $\sigma_v^2 = 0.025$, var(u) $= 0.225$, $\sigma_u^2 = 0.619$

Four values of N and T are considered. In order to investigate the effect of different size of N, we allow N to vary among 10, 20, 50,and 100, while T is fixed at 10. Similarly, we allow T to vary among 10, 20, 50, and 100, while N is fixed at 10 to investigate the effect of different size of T. This is done for each of the three different values of γ^*.

For a given set of parameter values, in each replication we generate the (NTx1) vector v from $N(0, \sigma_v^2)$ and generate the (Nx1) vector u from u=|U|, $U \sim N(0, \sigma_u^2)$. Then, the NTx1 vector y can be calculated from the model: $y_{it} = \alpha + v_{it} - u_i$, and the point estimates and confidence intervals are calculated from the data. We consider two estimation techniques: MLE and the FE method. MLE calculates the estimates of α, σ_v^2

and σ_u^2 by maximizing the half- normal likelihood function, and the point estimates and confidence intervals for u_i are calculated conditional on the estimated value of α, σ_v^2 and σ_u^2. MLE is based on a correct distributional assumption and might be expected to perform better than other techniques.

In the fixed effect method, the N individual effects $\alpha_1,.....\alpha_N$ and σ_v^2 are estimated first, and the overall constant term and firm specific inefficiencies are estimated by $\hat{\alpha} = \max_{j=1,...,N} \hat{\alpha}_j$ and $\hat{u}_i^* = \max_{j=1,...,N} \hat{\alpha}_j - \hat{\alpha}_i$. For the bootstrapping intervals, the percentile intervals and bias corrected and accelerated (BC_a) intervals are calculated. The iterated bootstrap is not considered due to its computational complexity.

By repeating this process a large number (say, m) of times, we obtain the distribution of the estimates. We calculate the means and standard errors of the point estimates, and the coverage probability of the confidence intervals is estimated by the fraction of times coverage occurs. For the FE & MargCB method, we use m=500 due to the relatively simple computations required, while we use m=300 for the MLE & BC method, and the FE & Bootstrap method. While we obtain point estimates and confidence intervals for each of the N cross-sectional units, we just report our results averaged over these N units.

4. Results

4.1 Estimation of Frontier Parameters

We begin our discussion of results by looking at the performance of the two estimation techniques in estimating the frontier parameters α, σ_v^2, and σ_u^2. Tables 1 and 2 provide the bias and standard error of the estimates.

Consider first the MLE estimates. The asymptotic theory for the MLE holds as $N \to \infty$, and our results confirm the fact that the MLE performs better for large N. In particular the biases of the estimates disappear as N gets large, not as T gets large. It is noteworthy that $\hat{\alpha}$ is biased *downward*, and so is $\hat{\sigma}_u^2$. This result was also found by Olson et al. (1980), but they had no explanation for it. However, based on the analysis of specific replications within our experiments, it appears that the bias occurs because the exact truncation point in the distribution of u_i is hard to estimate with small N. The estimation process can be regarded as allocating y_{it} among constant term, v_{it}, and u_i. In fact, it is not hard to distinguish the inefficiency term from the normal error even with small N, because we assume a half normal distribution for u_i that has a very different shape from a normal error. But a problem exists in that the constant term can not be distinguished easily from the inefficiency terms with small N. With small N, especially when σ_u^2 is big, the minimal value of u_i is not close to zero. In this case, some portion of u_i is implicitly allocated to the constant term. As a result, u_i tend to be underestimated and the constant term tends to be underestimated as well (due to the negative sign in front of u_i). This explains why substantial biases in $\hat{\alpha}$ are found for the same parameter values as substantial biases in the estimate of var(u), since the same logic explains the downward bias and its decrease with N and increase with γ^* in the estimation of var(u).

The influence of γ^* on the estimation of the constant term depends on the size of N. With small N, the minimal value of u_i is not close to zero and it tends to be large with bigger σ_u^2. This explains why the bias of $\hat{\alpha}$ increases as γ^* increases, at least when N is small. With N big enough to accurately identify the location of the distribution of u, γ^* does not have a significant influence on the bias of $\hat{\alpha}$. The effect of T on the bias of $\hat{\alpha}$ is also closely related to the size of N and γ^*. When N is big enough, the bias of $\hat{\alpha}$ decreases with T. Otherwise, the effect of T on the bias of $\hat{\alpha}$ depends on the size of γ^*. The bias decreases as T increases for small values of γ^* while the opposite happens for large values of γ^*. Meanwhile, the standard error of $\hat{\alpha}$ decreases as either N or T increases, but the effect of N is much stronger.

Table 1 and 2 also provide the bias and standard error of the FE estimators. The bias in the FE estimation of the constant term depends on two factors. First, the bias comes from the fact that the sample minimum of u_i is not zero. Clearly $\max_{j=1,\ldots,N} \alpha_j = \alpha - \min_{j=1,\ldots,N} u_j$. We implicitly regard $\min_{j=1,\ldots,N} u_j$ as zero by defining $\hat{\alpha} = \max_{j=1,\ldots,N} \hat{\alpha}_j$. So, when $\min_{j=1,\ldots,N} u_j$ is greater than zero, it results in a downward bias in $\hat{\alpha}$. This bias goes away as $N \to \infty$ and as $\min_{j=1,\ldots,N} u_j \to 0$. For small N, it also depends on the value of σ_u^2, since $\min_{j=1,\ldots,N} u_j$ tends to be larger for the larger values of σ_u^2.

The second and more important source of bias is the fact that $\max_{j=1,\ldots,N} \hat{\alpha}_j$ is a biased estimator of $\max_{j=1,\ldots,N} \alpha_j$ in finite samples. For fixed N, $\max_{j=1,\ldots,N} \hat{\alpha}_j$ is a consistent estimate of $\max_{j=1,\ldots,N} \alpha_j$ as $T \to \infty$. However, for finite values of T, $\max_{j=1,\ldots,N} \hat{\alpha}_j$ is biased *upward* as an

estimator of $\max_{j=1,\ldots,N} \alpha_j$. This bias is large when several of the largest α_j's are close each other (i.e., there are near ties for $\max_{j=1,\ldots,N} \alpha_j$) and the estimation error of $\hat{\alpha}_i$ is large.

The first source of bias decreases as N increases while the second source of bias is large when T is small and $\hat{\sigma}_v^2$ is large. The second bias also tends to be large with large N, because near ties for $\max_{j=1,\ldots,N} \alpha_j$ become more likely, and because the bias induced by the "max" operator is larger when the maximization is over a larger set of $\hat{\alpha}_i$. The final size and sign of the bias depend on the relative size of each source of bias. For example, $\hat{\alpha}$ has a negative bias at N=10, T=10, and γ^*=0.5, due to the dominance of the first factor, but it changes to a positive bias as N increases and the bias due to the first factor disappears.

4.2 Estimation of Efficiencies: MLE & BC method

We now turn our attention to the estimation of efficiencies. Table 3 provides the bias of the estimated inefficiencies, and the coverage rate of the confidence intervals. These are both averages over values i=1,...,N. As is easily noticed, the bias of estimated inefficiency is just a reflection of the bias of $\hat{\alpha}$, as explained in the previous section. The 90% confidence intervals are calculated according to the BC method. Column (3) in Table 3 shows the average width of the confidence intervals. These confidence intervals are for r = exp(-u) and are just a transformation of the confidence intervals for u. The coverage rates of the intervals are the same whether they are expressed in terms of u or r, but since $0 \leq r \leq 1$ the width of the confidence interval for r is easier to understand than the width of the confidence interval for u. Columns (4) and (5) show the frequency that the

true efficiency falls below the lower bound and above the upper bound of the confidence interval. For valid intervals each of these should equal 0.05. Column (6) shows the empirical coverage probability of the 90% confidence interval. The sum of the numbers in columns (4), (5) and (6) is always one.

The sample distribution of the efficiency estimates is not centered on the true value due to the downward bias in the estimation of inefficiency (or the upward bias in the estimation of efficiency). Thus, the probability of the true efficiency level to be smaller than the 90% lower bound is much greater than 0.05. However, it is also true that the probability that the true value is greater than the upper bound is a little larger than 0.05. It appears that the confidence intervals are too narrow in general (even not accounting for the bias problem). This occurs because the confidence intervals according to the BC method fail to incorporate the uncertainty in the estimated frontier parameters. As a result of both of these problems, the coverage rate is less than 0.9, and is as low as 0.29 in the extreme case of N=10, T=100, and $\gamma^*=0.9$.

The main determinant of the accuracy of the confidence intervals is the size of N. When N is large (e.g., N=100) the coverage rate of the intervals is reasonably close to 0.90 (e.g., 0.884 for N=100, T=10, $\gamma^*=0.5$). This is so because both of the problems discussed in the previous paragraph go away as N increases.

At this point, we need to recall that we are assuming the true distribution is known. Given this fact, the general performance of MLE is perhaps not as good as might be hoped. A natural question is what happens if we do not use the correct distribution of u_i. Table 4 gives some results for the case that MLE is not based on the true distribution of the data-generating process. The assumed distribution is half-normal, but the true

distribution of u is $N(\mu, \sigma_u^2)$ truncated at zero, with μ=0.5 or μ=-2. A glance at the results reveals that the performance is very bad for μ=0.5. $\hat{\alpha}$ and the estimated inefficiency have much bigger bias than in the case with correct information on the distribution of u_i. Consequently, the coverage probability is very low.

For the case of μ=-2, it seems that the incorrect assumption does not distort the results much. Especially for small N, $\hat{\alpha}$ is less biased and the coverage rate is actually higher than in the case assuming the correct distribution. This is presumably because the values of u_i are more concentrated near zero than when the sample is from a half–normal distribution, so that the inferred truncation point is much closer to zero. Table 5 might help to understand this mechanism. It shows the average sample minimum of u_i for samples from the different distributions. For μ=-2, the average sample minimum is much smaller than when μ=0 or μ=0.5. This makes it less likely that α is underestimated. These results are certainly distribution-specific, but they indicate the potential for serious bias and incorrect coverage probabilities if MLE is based on an incorrect distributional assumption.

4.3 Estimation of Efficiencies : FE & MargCB or MCB method

Table 6 and Table 7 summarize the performance of the FE estimates of efficiencies. Table 6 changes N holding T constant, while Table 7 changes T holding N constant. In both cases different values of γ^* are considered. We first consider the bias of the point estimates $\hat{u}_i^*$. The bias in the estimation of the inefficiencies can be defined in two ways, depending on whether $\hat{u}_i^*$ is viewed as an estimate of u_i^* or u_i. Column (2) in

Table 6 and Table 7 shows the bias of the estimated inefficiency as an estimate of u_i. It reflects the two sources of bias discussed above and it is just the mirror image of the bias of $\hat{\alpha}$. Column (3) shows the bias of the estimated inefficiency as an estimate of the relative inefficiency defined by $u_i^* = u_i - \min_{j=1,\ldots,N} u_j$. When we determine the bias in this way, it removes the effect of $\min_{j=1,\ldots,N} u_j \neq 0$ and only reflects the bias from the second source, which is the "max" operation over i=1,...,N. Our focus should really be on the second type of bias since $\hat{u}_i^*$ is most naturally thought of as an estimate of u_i^*. Thus, the estimated inefficiency can not be directly compared with MLE estimates because MLE estimates the inefficiency u_i defined in the absolute (not relative) sense.

As discussed in the previous section, the upward bias induced by the "max" operation disappears as $T \to \infty$. Accordingly, the effect of T is most important in the estimation of inefficiencies in the Fixed Effect model. The bias of estimated inefficiency decreases as N decreases, since the u_i's tend to be less tied together with small N, and the "max" operation is over less values of i. The bias also decreases as γ^* increases, since the α_i are more distinct and are estimated more precisely. Columns (4)-(11) give the width and coverage rates of the MargCB and MCB confidence intervals. The intervals are *very* conservative-they are very wide and they have coverage rates in excess of 0.90.

In fact, the coverage probability is close to one for all cases. In other to understand the reason, it would help to see how the best set is determined. Column (12) in Table 6 and Table 7 shows the average percentage of firms included in the best set and column (13) shows the probability that the best set includes the true best firm. A high percentage of firms are included in the best set, which implies a high degree of

uncertainty about which is the most efficient firm. The results clearly show how conservative the best set is. It includes the most efficient firm with a probability of almost one, even though it is designed to include the most efficient firm with a probability of at least 0.95. The MCB and MargCB coverage probabilities would be correct if the identity of the best firm were known, and uncertainty about which is the best firm is reflected in the size of the best set and in the excess coverage rates of the confidence intervals.

4.4 Estimation of Efficiencies : FE and Bootstrap method

MCB and MargCB provide very reliable but excessively wide confidence intervals in general. The confidence statements would not be useful if they are too wide, even if they are very reliable. We consider the bootstrap method as a possible alternative. Table 8 reports the results for confidence statements from the bootstrap method. The first four columns show the accuracy of the confidence intervals by the percentile method and the next four columns show the accuracy of the confidence intervals by the bias-corrected and accelerated (BC_a) percentile method. The bootstrap intervals are much narrower than the marginal MCB intervals, but the question is whether they are reliable.

It was discussed in Chapter 2 that the bootstrap estimator of the distribution of $[\max_{j=1,\ldots,N} \hat{\alpha}_j - \max_{j=1,\ldots,N} \alpha_j]$ is consistent in T if there are no ties for $\max_{j=1,\ldots,N} \alpha_j$. This implies that the validity of the percentile intervals depends on large T. For small values of T, the percentile intervals are not centered on the true values due to the bias problem previously discussed. Since the bias is small when we have large T and γ^* and small N, the coverage probability reaches almost 0.9 for these cases, but it falls in cases where the bias is big. The width of the intervals decreases as T or γ^* increases just as in MargCB or MCB. But,

the influence of N on the width of intervals is the opposite of that in MargCB or MCB. The intervals get narrower with larger N, while the bias increases as N increases. This explains why the coverage probabilities of the percentile intervals falls rapidly as N increases.

The results in Table 8 indicate that the BC_a percentile intervals are better than the uncorrected percentile intervals. Like the uncorrected percentile intervals, they are more accurate when T and γ^* are large and when N is small. When T and γ^* are small or N is large, they are a very considerable improvement over the percentile intervals, even though they do not succeed entirely in yielding correct coverage rates.

The bias corrected confidence intervals are obtained by shifting the bootstrap distribution by approximately twice the estimated bias in the bootstrap stage. If on average $[\max_{j=1,\ldots,N} \hat{\alpha}_j^{(b)} - \max_{j=1,\ldots,N} \hat{\alpha}_j]$ were the same as $[\max_{j=1,\ldots,N} \hat{\alpha}_j - \max_{j=1,\ldots,N} \alpha_j]$, we would expect a properly centered interval with a coverage rate of 0.9 after the bias is corrected. In our model, however, only some part of the bias gets be corrected. Some evidence on this point is given in Table 9, which shows the average of $\max_{j=1,\ldots,N} \alpha_j$, $\max_{j=1,\ldots,N} \hat{\alpha}_j$, and $\max_{j=1,\ldots,N} \hat{\alpha}_j^{(b)}$ over different values of N, T, and γ^*. The fourth column in the table shows the average bias in the FE estimates of $\max_{j=1,\ldots,N} \alpha_j$ and the last column shows the average bias in the bootstrap estimates. We see that $[\max_{j=1,\ldots,N} \hat{\alpha}_j^{(b)} - \max_{j=1,\ldots,N} \hat{\alpha}_j]$ is always smaller than $[\max_{j=1,\ldots,N} \hat{\alpha}_j - \max_{j=1,\ldots,N} \alpha_j]$ and the difference is substantial when γ^* is small and N is large. As a result, the bias correction is incomplete especially when γ^* is small and N is large.

5. Concluding Remarks

In this chapter, we have analyzed the finite sample properties of two techniques for the estimation of efficiencies, where the main difference lies in whether the technique uses a distributional assumption in estimating inefficiencies. The MLE & BC method, which uses a distributional assumption in estimating inefficiencies, does well when N is large but surprisingly poorly when N is small. When N is small, there is a large downward bias in the estimated inefficiencies, and this results in improperly centered confidence intervals with low coverage rates. This problem is especially severe when T and γ^* are large because the confidence intervals tend to be narrow with large T and large γ^*, while the bias is determined mainly by the size of N. However, there appears to be a large gain in precision and reliability of inference from making a distributional assumption when N is large relative to T.

For the Fixed Effects model things are rather different. The FE model estimates inefficiency relative to the sample maximum. The use of the "max" operation results in an upward bias in the estimation of the frontier, and therefore in an upward bias in the estimation of inefficiency (or a downward bias in the estimation of efficiency), which is a bias in the opposite direction as for the MLE & BC method. The bias in the FE estimated inefficiencies decreases as T or γ^* increases and as N decreases, but it can be very substantial when N is large and T is small.

MCB and MargCB based on the FE estimates yield very conservative confidence intervals. They tend to be very wide and to have coverage probabilities that are much higher than the nominal coverage probability. While MargCB and MCB do not suffer

from a bias problem, they are very wide in the same circumstances that the FE point estimates are seriously biased.

Confidence intervals based on bootstrapping of the fixed effectes tend to be improperly centered due to the bias due to the "max" operation. As a result the bootstrapping confidence intervals are not very reliable in general, though they are relatively reliable when T and γ^* are large and N is small. Using the bias-corrected bootstrap helps in terms of the accuracy of the intervals. It performs relatively well overall even though the bias-correction is not complete. Like the MargCB and MCB intervals, it performs better with small N and large T and γ^*, but it provides much narrower intervals than MargCB or MCB. The bootstrap method may be a reasonable alternative to the MLE when N is small.

The most important limitation of these results is that they are for a simple model without regressors. With regressors included, there are a number of additional factors to consider, including the nature of the regressors and their correlation or noncorrelation with the inefficiency term.

Table 1

Bias and Standard Error of Estimated Constant Term, σ_v^2 and var(u)
(N Changes with T Held Constant)

T=10	MLE			FE	
	Constant	σ_v^2	var(u)	Constant	σ_v^2
	bias (std.error) (1)	bias (std.error) (2)	bias (std.error) (3)	bias (std.error) (4)	bias (std.error) (5)
(γ^* =0.1)					
N=10	-.039 (.111)	-.003 (.030)	-.003 (.022)	.109 (.111)	-.002 (.032)
N=20	-.016 (.080)	-.002 (.022)	-.001 (.016)	.170 (.094)	-.001 (.022)
N=50	-.010 (.048)	-.001 (.014)	-.001 (.010)	.228 (.081)	-.001 (.015)
N=100	-.002 (.033)	-.001 (.010)	.000 (.007)	.271 (.070)	-.001 (.010)
(γ^* =0.5)					
N=10	-.042 (.112)	-.002 (.017)	- .010 (.068)	-.016 (.105)	-.001 (.018)
N=20	-.019 (.074)	-.001 (.012)	-.006 (.050)	.050 (.086)	-.001 (.012)
N=50	-.010 (.047)	-.001 (.008)	.000 (.030)	.107 (.070)	-.001 (.008)
N=100	-.005 (.032)	-.001 (.005)	.000 (.021)	.148 (.064)	.000 (.006)
(γ^* =0.9)					
N=10	-.049 (.095)	.000 (.003)	-.028 (.097)	-.074 (.087)	.000 (.004)
N=20	-.020 (.059)	.000 (.002)	-.013 (.070)	-.029 (.063)	.000 (.002)
N=50	-.008 (.034)	.000 (.002)	-.006 (.048)	.015 (.040)	.000 (.002)
N=100	-.003 (.024)	.000 (.001)	-.003 (.033)	.038 (.034)	.000 (.001)

Table 2

Bias and Standard Error of Estimated Constant Term, σ_v^2 and var(u)
(T Changes with N Held Constant)

N=10	MLE			FE	
	Constant	σ_v^2	var(u)	Constant	σ_v^2
	bias (std.error) (1)	bias (std.error) (2)	bias (std.error) (3)	bias (std.error) (4)	bias (std.error) (5)
(γ^* =0.1)					
T=10	-.039 (.111)	-.003 (.030)	-.003 (.022)	.109 (.111)	-.002 (.032)
T=20	-.028 (.086)	-.002 (.021)	-.004 (.016)	.050 (.080)	-.001 (.022)
T=50	-.023 (.055)	-.001 (.013)	-.004 (.013)	.008 (.056)	-.001 (.014)
T=100	-.014 (.047)	-.001 (.009)	-.003 (.013)	-.005 (.048)	-.001 (.010)
(γ^* =0.5)					
T=10	-.042 (.112)	-.002 (.017)	-.010 (.068)	-.016 (.105)	-.001 (.018)
T=20	-.039 (.095)	-.001 (.012)	-.010 (.060)	-.037 (.089)	-.001 (.012)
T=50	-.039 (.078)	-.001 (.008)	-.010 (.057)	-.055 (.076)	.000 (.008)
T=100	-.036 (.077)	-.001 (.005)	-.009 (.057)	-.056 (.073)	.000 (.005)
(γ^* =0.9)					
T=10	-.049 (.095)	.000 (.003)	-.028 (.097)	-.074 (.087)	.000 (.004)
T=20	-.055 (.089)	.000 (.002)	-.031 (.094)	-.077 (.083)	.000 (.002)
T=50	-.063 (.084)	.000 (.002)	-.034 (.092)	-.083 (.079)	.000 (.002)
T=100	-.067 (.082)	.000 (.001)	-.036 (.091)	-.084 (.077)	.000 (.001)

Table 3

Estimation of Efficiencies : MLE & BC method

	Bias		Confidence Interval (90%) for Efficiency (r)			
	constant	inefficiency (u)	width	prob(<LB)	prob(>UB)	coverage rate
	(1)	(2)	(3)	(4)	(5)	(6)
(γ^* =0.1, T=10)						
N=10	-.039	-.038	.203	.280	.056	.664
N=20	-.016	-.017	.235	.159	.052	.789
N=50	-.010	-.009	.254	.086	.049	.865
N=100	-.002	-.001	.260	.064	.051	.885
(γ^* =0.5, T=10)						
N=10	-.042	-.042	.209	.190	.066	.744
N=20	-.019	-.020	.210	.114	.059	.827
N=50	-.010	-.009	.210	.076	.052	.872
N=100	-.005	-.005	.210	.064	.051	.885
(γ^* =0.9, T=10)						
N=10	-.049	-.049	.095	.334	.077	.589
N=20	-.020	-.020	.093	.204	.079	.717
N=50	-.008	-.008	.092	.109	.067	.824
N=100	-.003	-.003	.092	.077	.062	.861
(γ^* =0.1, N=10)						
T=10	-.039	-.038	.203	.280	.056	.664
T=20	-.028	-.028	.186	.218	.060	.722
T=50	-.023	-.022	.147	.166	.053	.781
T=100	-.014	-.014	.112	.174	.070	.756
(γ^* =0.5, N=10)						
T=10	-.042	-.042	.209	.190	.066	.744
T=20	-.039	-.039	.156	.226	.065	.709
T=50	-.039	-.038	.104	.280	.068	.652
T=100	-.036	-.036	.075	.344	.106	.550
(γ^* =0.9, N=10)						
T=10	-.049	-.049	.095	.334	.077	.589
T=20	-.055	-.055	.069	.422	.081	.497
T=50	-.063	-.062	.045	.548	.066	.386
T=100	-.067	-.067	.032	.629	.077	.294

Table 4

Effect of Wrong Distribution on Estimation of Efficiencies: MLE & BC method
(Assumed Distribution is Half- Normal, True Distribution is $N(\mu, \sigma_u^2)$ truncated at zero)

	Bias				Confidence Int.(90%) for Efficiency (r)			
	constant	σ_v^2	var(u)	u	width	prob(<LB)	prob(>UB)	coverage rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(μ=0.5)								
γ^* =0.1								
N=10, T=10	-.335	-.002	-.003	-.334	.201	.831	.000	.169
N=50, T=10	-.307	.001	-.002	-.306	.254	.841	.000	.159
N=10, T=50	-.286	-.001	.002	-.285	.147	.955	.000	.045
γ^* =0.5								
N=10, T=10	-.142	-.001	-.002	-.141	.205	.417	.021	.562
N=50, T=10	-.101	.000	.012	-.100	.205	.278	.012	.710
N=10, T=50	-.107	-.001	.010	-.106	.098	.542	.034	.424
γ^* =0.9								
N=10, T=10	-.084	.000	.003	-.084	.090	.438	.054	.508
N=50, T=10	-.038	.000	.029	-.038	.087	.243	.031	.726
N=10, T=50	-.094	.000	-.002	-.094	.042	.636	.056	.308
(μ=-2)								
γ^* =0.1								
N=10, T=10	-.012	-.004	-.004	-.011	.187	.279	.082	.639
N=50, T=10	.028	-.002	-.001	.029	.253	.047	.098	.855
N=10, T=50	.002	-.001	-.006	.003	.144	.105	.088	.807
γ^* =0.5								
N=10, T=10	-.002	-.002	-.029	-.002	.211	.122	.099	.779
N=50, T=10	.032	-.001	-.014	.032	.216	.033	.104	.863
N=10, T=50	-.009	-.001	-.032	-.009	.108	.153	.109	.738
γ^* =0.9								
N=10, T=10	-.020	.000	-.045	-.020	.099	.230	.110	.660
N=50, T=10	.014	.000	-.035	.014	.098	.049	.117	.834
N=10, T=50	-.036	.000	-.052	-.035	.047	.412	.108	.480

Table 5

Average Sample Minimum for Different Distribution

	N=10			N=50		
	γ^*=0.1	γ^*=0.5	γ^*=0.9	γ^*=0.1	γ^*=0.5	γ^*=0.9
$\mu = 0$	.029	.067	.085	.006	.015	.020
$\mu = 0.5$	.266	.137	.120	.154	.035	.031
$\mu = -2$	.019	.043	.065	.004	.009	.012

Table 6

Estimation of Efficiencies : FE & MarCB or MCB
(N Changes with T Held Constant)

	Bias			MargCB Conf. Int.(90%) for r^*				MCB Conf. Int. (90%) for r^*				Best Set	
	constant	inefficiency		width	prob(<LB)	prob(>UB)	coverage rate	width	prob(<LB)	prob(>UB)	coverage rate	firms in best set	prob ($\max\alpha_j \in S$)
T=10		$(\hat{u}_i^* - u_i)$	$(\hat{u}_i^* - u_i^*)$										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(γ^*=0.1)													
N=10	.109	.110	.138	.494	.001	.002	.997	.546	.000	.001	.999	80%	.98
N=20	.170	.169	.184	.529	.000	.000	1.000	.596	.000	.000	1.000	79%	.99
N=50	.228	.228	.234	.560	.000	.000	1.000	.639	.000	.000	1.000	78%	.99
N=100	.271	.271	.274	.579	.000	.000	1.000	.666	.000	.000	1.000	77%	.99
(γ^*=0.5)													
N=10	-.016	-.016	.052	.418	.007	.004	.989	.472	.001	.002	.997	50%	1.00
N=20	.050	.049	.086	.460	.002	.000	.998	.531	.000	.000	1.000	47%	1.00
N=50	.107	.107	.121	.490	.000	.000	1.000	.577	.000	.000	1.000	44%	1.00
N=100	.148	.148	.155	.506	.000	.000	1.000	.603	.000	.000	1.000	42%	1.00
(γ^*=0.9)													
N=10	-.074	-.074	.012	.183	.015	.016	.969	.216	.005	.006	.989	26%	1.00
N=20	-.029	-.029	.020	.213	.009	.010	.981	.260	.001	.004	.995	22%	1.00
N=50	.015	.015	.035	.240	.002	.002	.996	.300	.000	.001	.999	19%	1.00
N=100	.038	.038	.048	.251	.001	.000	.999	.320	.001	.000	1.000	18%	1.00

Table 7

Estimation of Efficiencies : FE & MargCB or MCB
(T Changes with N Held Constant)

	Bias			MargCB Conf. Int.(90%) for r^*				MCB Conf. Int. (90%) for r^*				Best Set	
	constant	inefficiency		width	prob(<LB)	prob(>UB)	coverage rate	width	prob(<LB)	prob(>UB)	coverage rate	firms in best set	prob ($\max\alpha_j \in S$)
N=10		$(\hat{u}_i^* - u_i)$	$(\hat{u}_i^* - u_i^*)$										
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(γ^*=0.1)													
T=10	.109	.110	.138	.494	.001	.002	.997	.546	.000	.001	.999	80%	.98
T=20	.050	.049	.077	.399	.002	.000	.998	.443	.000	.000	1.000	74%	.99
T=50	.008	.008	.036	.293	.004	.003	.993	.328	.001	.002	.997	60%	1.00
T=100	-.005	-.005	.023	.227	.008	.003	.989	.256	.002	.001	.997	49%	1.00
(γ^*=0.5)													
T=10	-.016	-.016	.052	.418	.007	.004	.989	.472	.001	.002	.997	50%	1.00
T=20	-.037	-.038	.029	.317	.006	.004	.990	.362	.001	.002	.997	40%	1.00
T=50	-.055	-.055	.012	.203	.011	.010	.979	.236	.003	.004	.993	30%	1.00
T=100	-.056	-.056	.011	.144	.012	.018	.970	.169	.003	.008	.989	24%	1.00
(γ^*=0.9)													
T=10	-.074	-.074	.012	.183	.015	.016	.969	.216	.005	.006	.989	26%	1.00
T=20	-.077	-.077	.008	.129	.017	.013	.970	.153	.004	.005	.991	22%	1.00
T=50	-.083	-.083	.002	.079	.014	.015	.971	.094	.005	.006	.989	18%	1.00
T=100	-.084	-.084	.001	.054	.018	.015	.967	.065	.007	.004	.989	16%	1.00

Table 8

Confidence Intervals for Efficiency Estimates : FE & Bootstrap

	Percentile Intervals (90%) for r^*				BC_a Percentile Int. (90%) for r^*			
	width	prob(<LB)	prob(>UB)	coverage rate	width	prob(<LB)	prob(>UB)	coverage rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(γ^* =0.1, T=10)								
N=10	.355	.003	.288	.709	.330	.014	.133	.853
N=20	.345	.001	.449	.551	.326	.012	.185	.803
N=50	.324	.000	.661	.339	.311	.011	.264	.725
N=100	.304	.000	.793	.207	.294	.006	.375	.619
(γ^* =0.5, T=10)								
N=10	.250	.016	.136	.847	.249	.050	.081	.869
N=20	.245	.003	.247	.750	.245	.036	.116	.848
N=50	.230	.001	.421	.578	.230	.031	.155	.814
N=100	.218	.000	.585	.415	.217	.019	.224	.757
(γ^* =0.9, T=10)								
N=10	.112	.029	.089	.882	.112	.060	.066	.874
N=20	.112	.017	.135	.848	.112	.063	.084	.853
N=50	.108	.004	.230	.766	.109	.038	.100	.862
N=100	.104	.001	.351	.648	.104	.034	.124	.842
(γ^* =0.1, N=10)								
T=10	.355	.003	.288	.709	.330	.014	.133	.853
T=20	.280	.001	.221	.777	.260	.029	.082	.889
T=50	.195	.006	.152	.842	.186	.033	.067	.900
T=100	.144	.014	.156	.830	.140	.046	.079	.875
(γ^* =0.5, N=10)								
T=10	.250	.016	.136	.847	.249	.050	.081	.869
T=20	.191	.014	.107	.879	.190	.050	.063	.887
T=50	.128	.022	.074	.903	.127	.054	.048	.898
T=100	.093	.025	.094	.880	.093	.041	.075	.884
(γ^* =0.9, N=10)								
T=10	.112	.029	.089	.882	.112	.060	.066	.874
T=20	.084	.038	.074	.889	.084	.055	.062	.883
T=50	.055	.027	.057	.915	.055	.047	.047	.906
T=100	.039	.041	.058	.901	.039	.048	.052	.900

Table 9

Bias Correction in the Bootstrap Intervals

		$\max_{j=1,...,N} \alpha_j$ (1)	$\max_{j=1,...,N} \hat{\alpha}_j$ (2)	$\max_{j=1,...,N} \hat{\alpha}_j^{(b)}$ (3)	(2)-(1)	(3)-(2)
γ^*=0.1	N=10, T=10	.971	1.109	1.179	.138	.070
	N=10, T=50	.971	1.005	1.033	.034	.028
	N=50, T=10	.993	1.223	1.336	.230	.113
γ^*=0.5	N=10, T=10	.931	.981	1.018	.050	.037
	N=10, T=50	.931	.942	.953	.011	.011
	N=50, T=10	.985	1.107	1.173	.122	.067
γ^*=0.9	N=10, T=10	.910	.921	.930	.011	.009
	N=10, T=50	.910	.912	.914	.002	.002
	N=50, T=10	.980	1.016	1.036	.036	.021

Chapter 4

MARGINAL COMPARISONS WITH THE BEST AND THE EFFICIENCY MEASUREMENT PROBLEM

1. Introduction

Suppose that we have data on each of a set of N populations, indexed by a parameter θ_i, i = 1, 2, ..., N. The parameterization is such that a larger value of θ_i is "better" than a smaller value. For example, θ_i could be the mean survival time after medical treatment i; or, in the efficiency measurement example, more efficient firms have larger θ_i. Suppose that we order the θ_i as follows:

$$\theta_{(1)} \leq \theta_{(2)} \leq \ldots \leq \theta_{(N)}, \tag{4.1}$$

so that population (N) is best. The identity of the best population is *not* assumed to be known, which is the challenging aspect of the problem. Let $\theta \equiv (\theta_1, \theta_2, \ldots, \theta_N)$ be the (unordered) vector of parameters. We presume that we have data on each of the populations and correspondingly there is an estimate $\hat{\theta}$ of θ. Based on this estimate we wish to say which populations might be best, and to construct confidence intervals for the differences $\theta_{(N)} - \theta_i$, which measure the amount by which a given population differs from the best.

One solution to this problem is given by the technique of *multiple comparisons with the best*, or MCB. MCB constructs a set S of possibly best populations, and a set of intervals (L_i, U_i), such that:

(4.2) $\quad P[\,(N) \in S \text{ and } L_i \le \theta_{(N)} - \theta_i \le U_i \text{ for all } i\,] \ge 1-\alpha\,,$

where $1-\alpha$ is a chosen confidence level (e.g. 0.90). Thus with a given confidence level we have a set of populations that includes the best, and joint confidence intervals for all differences from the best.

An alternative to the multiple confidence intervals in (4.2) is a marginal (i.e., univariate) confidence interval for $\theta_{(N)} - \theta_i$, for a single given value of i. In the efficiency measurement example, this would amount to a confidence interval for the technical inefficiency of a given firm, which is a natural and useful object of interest. Perhaps surprisingly, the construction of a marginal confidence interval for $\theta_{(N)} - \theta_i$ is a previously unsolved problem. In this chapter we show how to construct these marginal confidence intervals. More precisely, for a given value of i, we provide a set S and and interval (L_i^m, U_i^m) such that $P[(N)\in S \textit{ and } L_i^m \le \theta_{(N)} - \theta_i \le U_i^m\,] \ge 1-\alpha$. The point is that marginal confidence intervals are natural to consider, and also that we would expect marginal confidence intervals to be narrower than joint ones.

2. Marginal Comparisons under Standard Assumptions

Throughout this chapter we will maintain the following two assumptions. First, we have an estimate $\hat{\theta}$, distributed as $N(\theta,\sigma^2 C)$ with C known. Second, either σ^2 is known, or we have an estimate $\hat{\sigma}^2$, independent of $\hat{\theta}$, such that $\hat{\sigma}^2/\sigma^2$ is distributed as χ_v^2/v. In any applications we envision, there will be enough degrees of freedom (v will be large enough) that we can effectively take σ^2 as known.

Standard MCB proceeds under the further assumption, which we will maintain in this section, that $C = kI_N$ with k known. This assumption is usually motivated by discussion of the "balanced one way model" (e.g., Hsu (1996), p. 43) in which we have independent observations y_{it} (i = 1,...,N, t = 1,...,T) distributed as $N(\theta_i, \sigma^2)$. In this case k = 1/T. This is also the case of the panel data regression model with fixed individual effects, if we treat the slope coefficients as known, as will be discussed below.

We define the following notation. E(½) is the (N-1) × (N-1) correlation matrix with all correlations equal to ½ (i.e., diagonal elements equal one, off-diagonal elements equal ½). Let z be a multivariate random variable distributed as student-t with dimension N-1, degrees of freedom v, and correlation matrix E(½). Define $d^*(\alpha)$ as the α-level critical value of $\max_{i=1,..,N-1} |z_i|$; i.e., $P[\max_{i=1,..,N-1} |z_i| \le d^*(\alpha)] = 1-\alpha$. Tabulations of $d^*(\alpha)$ can be found in Hsu (1996) or Horrace (1998). Define $h(\alpha) = d^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$, and define the set $S(\alpha) = \{i \mid \hat{\theta}_i \ge \max_{j=1,...,N} \hat{\theta}_j - h(\alpha)\}$. Define L_i and U_i as follows:

$$(4.3) \qquad L_i = \max[0, \min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - h(\alpha)] \,, \; U_i = \max[0, \max_{j \ne i} \hat{\theta}_j - \hat{\theta}_i + h(\alpha)]$$

Then MCB provides the statement (4.2) above, with $S = S(\alpha)$.

Our marginal comparison with the best is given by the following theorem.

THEOREM 1: Let $t^*(\alpha)$ be the two-sided α-level critical value of the (univariate) student-t distribution with v degrees of freedom; i.e., if z is distributed as student-t with v degrees of freedom, then $P[|z| \le t^*(\alpha)] = 1-\alpha$. Define $g(\alpha) = t^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$. Define the set $S(\alpha)$ as above. Define L_i^m and U_i^m as follows:

(4.4) $$L_i^m = \max[0, \min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - g(\alpha/2)],\ U_i^m = \max[0, \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha/2)].$$

Then

(4.5) $$P[\ (N) \in S(\alpha) \text{ and } L_i^m \leq \theta_{(N)} - \theta_i \leq U_i^m\ ] \geq 1-\alpha.$$

The proof is given in the Appendix. It is a relatively straightforward application of the Bonferroni inequality.

We can note that the marginal comparison (4.5) uses the $\alpha/2$-level critical value of a univariate student-t while the multiple comparison (4.2) uses the α-level critical value of an N-1 dimensional student t. It is not clear whether there is a general inequality between these, but for commonly chosen values of α (e.g. 0.05 or 0.10) the marginal intervals are narrower than the multiple intervals. For example, for $\alpha = 0.05$ and $v=\infty$, the univariate $\alpha/2$-level critical value is 2.11, while the N-1 variate α-level critical value is 2.34 for N=3, 3.16 for N=10, and so forth.

We may also wish to consider one-sided confidence intervals. One of the possible motivations for doing so is the following. In many applications, the lower bound for $\theta_{(N)} - \theta_i$ turns out to be zero for many observations, because the set S of possibly best populations is large. We might choose to forgo the calculation of a lower bound, in which case a tighter upper bound is possible. This result is given in the following theorem.

THEOREM 2: Let $g(\alpha)$ be defined as in the statement of Theorem 1. Then the following are true:

(4.6) $$P[\ (N) \in S(\alpha) \text{ and } \theta_{(N)} - \theta_i \le \max[0,\ \max_{j \ne i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha)]\ ] \ge 1-\alpha$$

(4.7) $$P[\theta_{(N)} - \theta_i \le \max[0,\ \max_{j \ne i} \hat{\theta}_j - \hat{\theta}_i + g(2\alpha)]] \ge 1-\alpha\ .$$

The proof is given in the Appendix, but we can note the following. Comparing (4.6) to (4.5), the fact that we make only one statement instead of two allows us to use the $\alpha/2$-level *one-sided* univariate student-t critical value, which is the same as the α-level two-sided critical value, instead of the $\alpha/2$-level two-sided critical value. (For example, for $\alpha = 0.05$ and $v=\infty$, we use 1.96 instead of 2.24.) Considering (4.7), we note that the upper bound does not require the definition of the possibly best set $S(\alpha)$. If we do not wish to consider $S(\alpha)$, we can devote the full confidence level $1-\alpha$ to the upper bound, and we can use the α-level one-sided critical value of student-t, which is the same as the 2α-level two-sided critical value. (Thus, for example, with $\alpha = 0.05$ and $v=\infty$, we can now use the critical value 1.64 instead of 1.96 or 2.24.) As a result we get more a precise upper bound.

3. Marginal Comparison with General Covariance Structure

The previous section considered the commonly-assumed special case that the covariance matrix of $\hat{\theta}$ was proportional to an identity matrix. In this section we consider the general case that $\hat{\theta}$ is distributed as $N(\theta,\sigma^2 C)$ with C known but unrestricted. This case arises in, among other cases, the panel data regression model with nontrivial regressors.

We first need to define a little notation. For a given value of j, define δ_j as the $(N-1)\times 1$ vector whose typical element is of the form $\theta_i-\theta_j$, for $i = 1,...,N$, $i\neq j$. Formally $\delta_j=D_j\theta$ where D_j is an $(N-1)\times N$ differencing matrix. The covariance matrix of $\hat{\delta}_j$ is $\sigma^2 B_j$, where $B_j = D_j C D_j'$. Let R_j be the corresponding correlation matrix. In the special case that C is proportional to identity, $R_j = E(\frac{1}{2})$, as discussed in the previous section. In the general case, R_j will depend on j and has no special structure, but it is easily calculated. Define $d_j^*(\alpha)$ as the two-sided α-level critical value of the multivariate student-t distribution with dimension N-1, degrees of freedom v, and correlation matrix R_j. This critical value will typically depend on j and will generally need to be calculated numerically (e.g. by a simulation), since tabulation is impossible except in special cases.

Now define $\hat{\sigma}_{ij}^2 = \hat{\sigma}_{ji}^2 = [\hat{\sigma}^2(C_{ii}+C_{jj}-2C_{ij})]$, $h_{ji}(\alpha) = d_j^*(\alpha)\,\hat{\sigma}_{ij}$, $L_i^j = \hat{\theta}_j-\hat{\theta}_i - h_{ji}(\alpha)$, $U_i^j = \hat{\theta}_j-\hat{\theta}_i + h_{ji}(\alpha)$, the possibly best set $S(\alpha) = \{i \mid U_j^i \geq 0\ \forall\ j\neq i\ \} = \{i \mid \hat{\theta}_i \geq \hat{\theta}_j - h_{ij}(\alpha)\ \forall\ j\neq i\}$, and the lower and upper bounds $L_i = \max[0, \min_{j\in S(\alpha)} L_i^j]$ and $U_i = \max[0, \max_{j\neq i} U_i^j]$. Then MCB provides the statement (4.2) above, with $S = S(\alpha)$.

We now provide the corresponding marginal comparison result.

THEOREM 3: Define the set $S(\alpha)$ as above, and let $t^*(\alpha)$ be the two-sided α-level critical value of the univariate student-t distribution. Define $g_{ij}(\alpha)= t^*(\alpha)\,\hat{\sigma}_{ij}$. Define L_i^m and U_i^m as follows:

(4.8) $$L_i^m = \max[0, \min_{j\in S(\alpha)} (\hat{\theta}_j - g_{ij}(\alpha/2)) - \hat{\theta}_i],\quad U_i^m = \max[0, \max_{j\neq i} (\hat{\theta}_j + g_{ij}(\alpha/2)) - \hat{\theta}_i].$$

Then

(4.9) $$P[\ (N) \in S(\alpha) \text{ and } L_i^m \le \theta_{(N)} - \theta_i \le U_i^m\ ] \ge 1-\alpha.$$

The proof is similar to the proof of Theorem 1 and is therefore omitted.

As in the standard case, we may also consider one-sided confidence intervals. The following Theorem (also presented without proof) is the result corresponding to Theorem 2.

THEOREM 4: The following are true:

(4.10) $$P[\ (N) \in S(\alpha) \text{ and } \theta_{(N)} - \theta_i \le \max[0,\ \max_{j \ne i} (\hat{\theta}_j + g_{ij}(\alpha)) - \hat{\theta}_i]\ ] \ge 1-\alpha$$

(4.11) $$P[\ \theta_{(N)} - \theta_i \le \max[0,\ \max_{j \ne i} (\hat{\theta}_j + g_{ij}(2\alpha)) - \hat{\theta}_i]\ ] \ge 1-\alpha\ .$$

As in the previous section, one possible motivation for one-sided confidence intervals is that they yield more precise upper bounds. However, in the case of general covariance structure the one-sided intervals given in equation (4.11) also offer considerable computational advantages, because they do not require the calculation of the possibly best set $S(\alpha)$. The calculation of $S(\alpha)$ requires the N critical values $d_j^*(\alpha)$, $j = 1,\ldots,N$, each of which is from an N-1 dimensional student-t distribution and is generally calculable only numerically (via simulation). Especially when N is large, this is a very complicated and time-consuming set of calculations.

4. Application to the Efficiency Measurement Problem

In this section we consider some empirical applications of our marginal comparisons with the best to the efficiency measurement problem. Here we will give a very brief discussion of the problem. See Chapter 2 for more detail on the problem.

We begin with the fixed-effect panel data regression model:

$$(4.12) \qquad y_{it} = \alpha_i + x_{it}'\beta + v_{it}, \quad i = 1,...,N, \quad t = 1,...T.$$

Here i indexes firms and t indexes time periods. For purposes of exposition we assume a "balanced" panel (T is the same for all i). We assume that the v_{it} are iid $N(0, \sigma_v^2)$ and we treat the explanatory variables x_{it} as fixed. The parameters of interest are the intercepts α_i, which correspond to the θ_i in the previous sections. In our applications the estimator of β, say, $\hat{\beta}$ is the fixed-effect or "within" estimator obtained by least squares of $(y_{it} - \bar{y}_i)$ on $(x_{it} - \bar{x}_i)$, where $\bar{y}_i$ and $\bar{x}_i$ are the means of the T observations for firm i. Then we obtain an estimate of α_i as $\hat{\alpha}_i = \bar{y}_i - \bar{x}_i'\hat{\beta}$.

Here the specific context is the efficiency measurement problem, in which y is the logarithm of output and x is a vector of functions of inputs into the productive process. Larger α_i is better because it corresponds to more output for the same inputs. We define $u_i^* = \alpha_{(N)} - \alpha_i$, where $\alpha_{(N)}$ is the largest of the N α's, and the *technical efficiency* of firm i is typically defined as $r_i^* = \exp(-u_i^*)$. Since $u_i^* \geq 0$, $0 \leq r_i^* \leq 1$. MCB and our marginal comparisons with the best procedures will provide confidence intervals for u_i^*, and these are easily converted into confidence intervals for r_i^*. In particular, if (with a given

probability) $L \leq u_i^* \leq U$, then $\exp(-U) \leq r_i^* \leq \exp(-L)$, so that lower bounds for u_i^* convert to upper bounds for r_i^* and conversely.

Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_N)$. Under our assumptions the covariance matrix of $\hat{\alpha}$ is $V(\hat{\alpha}) = (\sigma_v^2/T)I_N + \bar{x}\, V(\hat{\beta})\, \bar{x}'$, where $V(\hat{\beta})$ is the covariance matrix of $\hat{\beta}$ and $\bar{x}$ is the matrix whose i^{th} row is $\bar{x}_i'$. In general this is not proportional to an identity matrix and so we should allow for a general covariance structure. The methods of section 2 ("standard" MCB or marginal comparisons) would apply if β were known, and can be viewed as applying approximately if the proportion of $V(\hat{\alpha})$ due to the variance of $\hat{\beta}$ is small. This may generally be so when N is large relative to T, as discussed in Horrace and Schmidt (1998). However, we present only the results that allow for the general covariance structure.

5. Empirical Applications

5.1 Indonesian Rice Farms

These techniques are applied to the same three data sets as were used in Chapter 2. We first analyze the data of Erwidodo (1990), which contain information on N = 171 rice farms for T = 6 growing seasons. Our results are given in Table 1. We choose α = 0.10 (hence 90% confidence intervals). We give results for the three most efficient (best) firms, the firms at the 75th percentile, 50th percentile, 25th percentile, and the two least efficient (worst) firms. For each firm, we present the value of $\hat{\alpha}_i$; the point estimate of technical efficiency; the confidence intervals corresponding to the marginal and multiple comparisons with the best; and the corresponding one-sided marginal and one-sided

multiple comparisons with the best. The one-sided marginal comparisons correspond to equation (4.11) of section 3.

In these data we do not estimate the intercepts α_i very precisely. This occurs because we have only six observations per firm and because the value of σ_v^2 turns out to be large relative to the variation in the α_i. Correspondingly, our confidence intervals are rather wide. In fact, they are wide enough to suggest that the efficiency measurement exercise has more or less failed to distinguish efficient and inefficient firms. The possibly best set S contains 98 of the 171 firms, and the other 73 firms are sufficiently close to being in the possibly best set that the MCB upper bounds for r_i^* equal one for all firms, even for the two-sided intervals. (For the one-sided intervals, the upper bound is automatically one, but this is an identity, not a data-determined outcome.)

The marginal comparison intervals are considerably shorter than the MCB intervals, as they should be. They use the 5% critical value of the univariate student-t distribution, 1.96, while the multiple intervals use 10% critical values of the 170-dimensional student-t distribution, which vary a little over comparison populations (i.e. d_j^* above depends on j) but equal 3.18 on average. The greater precision of the marginal as opposed to multiple confidence intervals is most noticeable for the more efficient firms. For example, for the most efficient firm compare the marginal interval of [0.74,1] to the multiple interval of [0.58,1]. The extra width of the MCB intervals is the price one has to pay for making a multiple statement, of course (the so-called "multiplicity effect").

5.2 Texas Utilities

We next analyze the data of Kumbhakar (1994). The results are given in Table 2. The confidence intervals (both marginal and multiple) are much narrower for this data set than for the previous one. We are now able to make statements about efficiencies that are precise enough to be meaningful. For example, only two observations are in the possibly best set, and the confidence interval for the efficiency of the (apparently) most efficient firm is [0.95,1]. This occurs primarily because σ_v^2 is smaller and T is larger here, so that the α_i are estimated more precisely. In the case of MCB, the intervals will also tend to be narrower because the multiplicity effect is not as strong with N = 10 as it is with N = 171. The marginal intervals are narrower than the multiple intervals but the difference in width is not as large as it was in the previous data set, again because the multiplicity effect is weaker here. Numerically, the marginal intervals use the univariate student-t 5% critical value of 1.96, while the multiple intervals use 10% nine-variate student-t critical values, which are on average equal to 2.38.

5.3 Egyptian Tileries

Lastly, we analyze the data of Seale (1990). The results are given in Table 3. In terms of the width of the confidence intervals, the results are somewhere between those for the Indonesian rice farms and those for the Texas utilities. This is true also in terms of the extent to which the marginal intervals are narrower than the multiple intervals. The marginal intervals use the critical value 1.96, while the multiple intervals use critical values that are on average equal to 2.70. As was the case in the Indonesian rice farm data, the difference between the marginal and multiple intervals is considerable for the more efficient firms but is not very large for the less efficient firms.

6. Concluding Remarks

In this chapter we have considered the general problem of creating confidence intervals for measures of the difference between a given population and the best population. More precisely, population i is characterized by a parameter θ_i, and we wish to construct a confidence interval for the difference $\theta_{(N)} - \theta_i$, where $\theta_{(N)} = \max_{j=1,\ldots,N} \theta_j$. This is a challenging problem because we do not know which population is best. One solution is given by MCB, which provides the complete set of N such confidence intervals, all of which hold simultaneously with at least a specified confidence level. Perhaps surprisingly, the seemingly simpler problem of providing a confidence interval for a single difference $\theta_{(N)} - \theta_i$ had not previously been solved. We provide these confidence intervals, and refer to them as *marginal* comparisons with the best.

Whether one prefers multiple or marginal comparisons will no doubt depend on the context. For an example of the arguments in favor of multiple comparisons, see Hsu (1996, p. 7). However, in some cases a marginal comparison may be natural. It seems reasonable to be able to perform either type of inference, just as one wishes to be able to test a set of hypotheses either individually or jointly.

In the context of the efficiency measurement problem, marginal comparisons correspond to the construction of the confidence interval for a given firm's technical efficiency level, and this is indeed a natural thing to consider. For example, models that assume a distribution for u_i yield marginal confidence intervals, constructed in somewhat more straightforward ways than here. See Horrace and Schmidt (1996) or Koop *et al.* (1997) for some examples. A marginal comparison with the best is directly comparable,

and provides evidence on the gain in precision from assuming a distribution for u_i. These comparisons are harder when MCB is used because the multiplicity effect and the effect of assuming or not assuming a distribution become confounded.

APPENDIX

Proof of Theorem 1

As in the text, suppose that z is a multivariate random variable distributed as student-t with dimension N-1, degrees of freedom v, and correlation matrix E(½). Define $d_1^*(\alpha)$ as the α-level critical value of $\max_{j=1,..,N-1} z_i$; i.e., $P[\max_i z_i \le d_1^*(\alpha)] = 1-\alpha$. Note that $d_1^*(\alpha)$ is the one-sided critical value corresponding to the two-sided critical value $d^*(\alpha)$ used in MCB, and that $d_1^*(\alpha/2) = d^*(\alpha)$. Similarly define $h_1(\alpha) = d_1^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$ and note that $h_1(\alpha/2) = h(\alpha)$, with $h(\alpha) = d^*(\alpha)(2k\hat{\sigma}^2)^{1/2}$, as used in MCB.

Consider the event $E_i(\alpha)=\{\theta_{(N)} - \theta_j \le \hat{\theta}_{(N)} - \hat{\theta}_j + h_1(\alpha)\ \forall j\neq(N)\}$. This is the one-sided multiple comparisons with a control (MCC) event, with (N) as control, and is constructed so that $P[E_1(\alpha)] = 1-\alpha$. See Dunnett (1955, 1964) or the discussion in Hsu (1996, chapter 3). The event $E_1(\alpha)$ implies the event $\{(N) \in S_1(\alpha)\}$, where $S_1(\alpha)$ is the set of indices $S_1(\alpha) = \{i \mid \hat{\theta}_i \ge \max_{j=1,..,N} \hat{\theta}_j - h_1(\alpha)\}$. Note that $S_1(\alpha/2) = S(\alpha) \subseteq S(\alpha/2)$. Therefore

(A1) $$P[(N) \in S_1(\alpha/2)] = P[(N) \in S(\alpha)] \ge 1-\alpha/2,$$

a standard result of the "ranking and selection" literature; e.g., see Gupta (1965).

Now pick a value of i (= 1,...,N), and consider the event $A_i(\alpha) = \{\hat{\theta}_{(N)} - \hat{\theta}_i - g(\alpha) \le \theta_{(N)} - \theta_i \le \hat{\theta}_{(N)} - \hat{\theta}_i + g(\alpha)\}$, where $g(\alpha)$ was defined in the statement of the Theorem. Note that $g(\alpha)$ was constructed so that $P[A_i(\alpha)] = 1-\alpha$. By the Bonferroni inequality, it follows from (A1) and $P[A_i(\alpha/2)] = 1-\alpha/2$ that

(A2) $P[\ (N) \in S(\alpha) \text{ and } A_i(\alpha/2)\ ] \geq 1-\alpha$.

This inequality is not immediately useful because it is not in terms of observable quantities, since (N) is unknown. So, we need to show that the event $\{(N) \in S(\alpha)$ and $A_i(\alpha/2)\}$ implies the marginal comparison event given in (4.5) of the main text. Consider first the lower bound. The event whose probability is given in (A2) implies that

(A3) $$\min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - g(\alpha/2) \leq \hat{\theta}_{(N)} - \hat{\theta}_i - g(\alpha/2) \leq \theta_{(N)} - \theta_i.$$

Also $0 \leq \theta_{(N)} - \theta_i$. Thus the event whose probability is given in (A2) implies the lower bound

(A4) $$\max[0, \min_{j \in S(\alpha)} \hat{\theta}_j - \hat{\theta}_i - g(\alpha/2)] \leq \theta_{(N)} - \theta_i.$$

The treatment of the upper bound is similar. If (N) = i, then $\theta_{(N)} - \theta_i = 0$. If $(N) \neq i$, the event in (A2) implies

(A5) $$\theta_{(N)} - \theta_i \leq \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha/2).$$

Therefore we have the upper bound

(A6) $$\theta_{(N)} - \theta_i \leq \max[0, \max_{j \neq i} \hat{\theta}_j - \hat{\theta}_i + g(\alpha/2)].$$

Finally, since the event $\{(N) \in S(\alpha)\}$ and the bounds (A4) and (A6) are implied by the event in (A2), they hold with at least the probability of that event; that is, with a probability no smaller than $1-\alpha$.

Proof of Theorem 2

As in the proof of Theorem 6, we have $P[(N) \in S(\alpha)] \geq 1-\alpha/2$. Now we also have $P[\theta_{(N)} - \theta_i \leq \hat{\theta}_{(N)} - \hat{\theta}_i + g(\alpha)] = 1-\alpha/2$, since $g(\alpha)$ is based on the α-level two-sided student-t critical value, or equivalently the $\alpha/2$-level one-sided critical value. Thus the Bonferroni inequality implies that $P[(N) \in S(\alpha)$ and $\theta_{(N)} - \theta_i \leq \hat{\theta}_{(N)} - \hat{\theta}_i + g(\alpha)] \geq 1-\alpha$. Then the same logic as was used in the discussion leading up to (A6) yields the result in (4.6).

To establish equation (4.7), we note that the upper bound does not require the definition of the possibly best set $S(\alpha)$. We simply start with the statement: $P[\theta_{(N)} - \theta_i \leq \hat{\theta}_{(N)} - \hat{\theta}_i + g(2\alpha)] = 1-\alpha$, which follows from the fact that the two-sided 2α-level critical value in $g(2\alpha)$ is the same as the α-level one-sided critical value. Then we again apply the same logic as was used in the discussion leading up to (A6) to obtain (4.7).

Table 1

90% Confidence Intervals for Technical Efficiency

Indonesian Rice Farms

Firm No.	$\hat{\alpha}_i$	$\hat{u}_i^*$	MargCB		MCB		1-Sided MargCB		1-Sided MCB	
			LB	UB	LB	UB	LB	UB	LB	UB
164	5.556	1	.737	1	.583	1	.840	1	.615	1
118	5.486	.932	.642	1	.508	1	.730	1	.534	1
163	5.484	.930	.643	1	.509	1	.729	1	.535	1
:	:	:	:	:	:	:	:	:	:	:
31	5.072	.616	.421	1	.328	1	.478	1	.346	1
:	:	:	:	:	:	:	:	:	:	:
15	4.966	.554	.379	1	.300	1	.432	1	.315	1
:	:	:	:	:	:	:	:	:	:	:
16	4.859	.498	.340	1	.266	1	.387	1	.280	1
:	:	:	:	:	:	:	:	:	:	:
117	4.586	.379	.259	.974	.203	1	.295	1	.214	1
45	4.550	.365	.250	.940	.197	1	.285	1	.208	1

Table 2

90% Confidence Intervals for Technical Efficiency

Texas Utilities

Firm No.	$\hat{\alpha}_i$	$\hat{u}_i^*$	MargCB		MCB		1-Sided MargCB		1-Sided MCB	
			LB	UB	LB	UB	LB	UB	LB	UB
5	-4.995	1	.960	1	.945	1	1	1	.967	1
3	-5.083	.916	.806	1	.797	1	.842	1	.815	1
10	-5.145	.861	.773	.979	.765	.984	.802	1	.780	1
1	-5.176	.835	.775	.976	.769	.984	.795	1	.780	1
8	-5.194	.820	.764	.964	.759	.973	.783	1	.769	1
9	-5.211	.806	.759	.961	.755	.972	.775	1	.763	1
2	-5.218	.801	.740	.934	.735	.942	.760	1	.745	1
7	-5.236	.786	.722	.913	.716	.920	.743	1	.728	1
6	-5.237	.786	.720	.910	.715	.917	.742	1	.726	1
4	-5.267	.762	.711	.901	.707	.910	.728	1	.716	1

Table 3

90% Confidence Intervals for Technical Efficiency

Egyptian Tileries

Firm No.	$\hat{\alpha}_i$	$\hat{u}_i^*$	MargCB		MCB		1-Sided MargCB		1-Sided MCB	
			LB	UB	LB	UB	LB	UB	LB	UB
14	.988	1	.805	1	.739	1	.868	1	.763	1
24	.983	.994	.790	1	.724	1	.859	1	.752	1
25	.977	.989	.789	1	.724	1	.856	1	.751	1
:	:	:	:	:	:	:	:	:	:	:
18	.940	.953	.737	1	.670	1	.810	1	.698	1
:	:	:	:	:	:	:	:	:	:	:
4	.877	.895	.708	1	.648	1	.772	1	.673	1
:	:	:	:	:	:	:	:	:	:	:
6	.549	.645	.518	.956	.478	1	.560	1	.495	1
:	:	:	:	:	:	:	:	:	:	:
7	.399	.555	.451	.816	.416	.885	.484	1	.429	1
8	.282	.493	.392	.744	.358	.815	.424	1	.371	1

BIBLIOGAPHY

Aigner, D.J., C.A.K. Lovell and P. Schmidt (1977), "Formulation and Estimation of Stochastic Frontier Production Functions," *Journal of Econometrics* 6, 21-37.

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

Broeck, J. van den, G. Koop, J. Osiewalski and M. Steel (1994), "Stochastic Frontier Models: A Bayesian Perspective," *Journal of Econometrics* 61, 273-303.

Coelli,T. (1995), "Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis," *Journal of Productivity Analysis* 6, 247-268.

Dorfman, J.H. (1997), *Bayesian Economics through Numerical Methods: A Guide to Econometrics and Decision-Making with Prior Information*, New York: Springer Verlag.

Dunnett, C.W. (1955), "A Multiple Comparison Procedure for Comparing Several Treatments with a Control," *Journal of the American Statistical Association* 50, 1096-1121.

Dunnett, C.W. (1964), "New Tables for Multiple Comparisons with a Control," *Biometrics* 20, 482-491.

Edwards, D.G. and J.C. Hsu (1983), "Multiple Comparisons with the Best Treatment," *Journal of the American Statistical Association* 78, 965-971. Corrigenda (1984), *Journal of the American Statistical Association* 79, 965.

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia: Society for Industrial and Applied Mathematics.

Efron, B. (1985), "Bootstrap Confidence Intervals for a Class of Parametric Problems," *Biometrika* 72, 45-58.

Efron, B. and R.J. Tibshirani (1993), *An Introduction to the Bootstrap*, New York: Chapman and Hall.

Erwidodo (1990), "Panel Data Analysis on Farm-Level Efficiency, Input Demand and Output Supply of Rice Farming in West Java, Indonesia," unpublished dissertation, Department of Agricultural Economics, Michigan State University.

Freedman, D. A. (1981), "Bootstrapping Regression Models," *The Annals of Statistics* 9,

No. 6, 1218-1228.

Gong, B. and R.C. Sickles (1992), "Finite Sample Evidence on the Performance of Stochastic Frontiers and Data Envelopment Analysis Using Panel Data," *Journal of Econometrics* 51, 259-284.

Greene, W.H. (1990), "A Gamma-Distributed Stochastic Frontier Model," *Journal of Econometrics* 46, 141-64.

Gupta, S.S. (1965), "On Some Multiple Decision (Selection and Ranking) Rules," *Technometrics* 7, 225-245.

Hall, P., W. Härdle and L. Simar (1993), "On the Inconsistency of Bootstrap Distribution Estimators," *Computational Statistics and Data Analysis* 16, 11-18.

Hall, P.,W. Härdle and L. Simar (1995), "Iterated Bootstrap with Applications to Frontier Models," *Journal of Productivity Analysis* 6, 63-76.

Hochberg, Y. and A.C. Tamhane (1987), *Multiple Comparison Procedures*, New York: John Wiley and Sons.

Horrace, W.C. (1998), "Tables of Percentages of the k-Variate Normal Distribution for Large Values of k," *Communications in Statistics: Simulation and Computation* 27, 823-831.

Horrace, W.C. and P. Schmidt (1996), "Confidence Statements for Efficiency Estimates from Stochastic Frontier Models," *Journal of Productivity Analysis* 7, 257-282.

Horrace, W.C. and P. Schmidt (1999), "Multiple Comparisons with the Best, with Economic Applications," *Journal of Applied Econometrics*, forthcoming.

Hsu, J.C. (1981), "Simultaneous Confidence Intervals for All Distances from the Best," *Annals of Statistics* 9, 1026-1034.

Hsu, J.C. (1984), "Constrained Simultaneous Confidence Intervals for Multiple Comparisons with the Best," *Annals of Statistics* 12, 1145-1150.

Hsu, J.C. (1996), *Multiple Comparisons: Theory and Methods*, London: Chapman and Hall.

Jondrow, J., C.A.K. Lovell, I. Materov and P. Schmidt (1982), "On the Estimation of Technical Inefficiency in the Stochastic Frontier Production Function Model," *Journal of Econometrics* 19, 233-238.

Koop, G., M.F. Steel and J. Osiewalski (1995), "Posterior Analysis of Stochastic Frontier Models using Gibbs Sampling," *Computational Statistics* 10, 353-373.

Koop, G., J. Osiewalski and M.F. Steel (1997), "Bayesian Efficiency Analysis through Individual Effects: Hospital Cost Frontiers," *Journal of Econometrics* 76, 77-106.

Kumbhakar, S.C. (1996), "Estimation of Cost Efficiency with Heteroscedasticity: An Application to Electric Utilities," *Journal of the Royal Statistical Society, Series D* 45, 319-335.

Lee, Y.H. (1991), "Panel Data Models with Multiplicative Individual and Time Effects: Applications to Compensation and Frontier Production Functions," unpublished dissertation, Department of Economics, Michigan State University.

Lee, Y.H. and P. Schmidt (1993), "A Production Frontier Model with Flexible Temporal Variation in Technical Efficiency, in *The Measurement of Productive Efficiency*, (Eds. H.O. Fried, C.A.K. Lovell and S.S. Schmidt), New York: Oxford University Press.

Olson, J. A., P. Schmidt and D. Waldman (1980), "A Monte Carlo Study of Estimators of Stochastic Frontier Production Functions," *Journal of Econometrics* 13, 67-82.

Osiewalski, J. and M. Steel (1998), "Numerical Tools for the Bayesian Analysis of Stochastic Frontier Models," *Journal of Productivity Analysis* 10, 103-117.

Park, B.U. and L. Simar (1994), "Efficient Semiparametric Estimation in a Stochastic Frontier Model," *Journal of the American Statistical Association* 89, 929-936.

Pitt, M.M. and L.F. Lee (1981), "The Measurement and Sources of Technical Inefficiency in the Indonesian Weaving Industry," *Journal of Development Economics* 9, 43-64.

Seale J. L. (1990), "Estimating Stochastic Frontier Systems with Unbalanced Panel Data: The Case of Floor Tile Manufactories in Egypt," *Journal of Applied Econometrics* 5, 59-79.

Schmidt P. and R.C. Sickles (1984), "Production Frontiers and Panel Data," *Journal of Business and Economic Statistics* 2, 367-74.

Simar, L. (1992), "Estimating Efficiencies from Frontier Models with Panel Data: A Comparison of Parametric, Non-Parametric and Semi-Parametric Methods with Bootstrapping," *Journal of Productivity Analysis* 3, 171-203.

Stevenson, R.E. (1980), "Likelihood Functions for Generalized Stochastic Frontier Estimation," *Journal of Econometrics* 13, 57-66.