ESTIMATION OF TECHNICAL EFFICIENCY FROM PARAMETRIC AND NON-PARAMETRIC PRODUCTION FRONTIERS

Evangelia Desli, Ph.D.

University of Connecticut, 1999

The dissertation consists of three essays that are developed independently. They all contribute to estimation of statistical production frontiers and technical efficiency using alternative techniques. The first essay relates to econometric estimation of production frontiers from panel data and modeling of technical efficiency and technical change. The second essay develops a methodology for estimation of a statistical production frontier using mathematical programming and bootstrapping techniques. Finally, the third essay estimates a non-parametric frontier using Data Envelopment Analysis and bootstrapping. Each essay includes a methodological extension with an empirical application.

ESTIMATION OF TECHNICAL EFFICIENCY FROM PARAMETRIC AND NON-PARAMETRIC PRODUCTION FRONTIERS

Evangelia Desli, Ph.D.

University of Connecticut, 1999

B.A., Aristotle University of Thessaloniki, 1990

M.A., University of Connecticut, 1996

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

1999

UMI Number: 9926243

UMI
300 North Zeeb Road
Ann Arbor, MI 48103

APPROVAL PAGE

Doctor of Philosophy Dissertation

ESTIMATION OF TECHNICAL EFFICIENCY FROM PARAMETRIC AND NON-PARAMETRIC PRODUCTION FRONTIERS

Presented by

Evangelia Desli, B.A., M.A.

Major Advisor ______________________ Subhash C. Ray

Associate Advisor ______________________ Kathleen Segerson

Associate Advisor ______________________ Dennis R. Heffley

University of Connecticut

1999

ACKNOWLEDGEMENTS

I am especially grateful to Prof. Subhash C. Ray, my Major Advisor, for his guidance, encouragement, friendship and support during the last five years. He has always showed concerned for me as an individual, both professionally and personally.

My sincere thanks is extended to Prof. Kathleen Segerson and Prof. Dennis R. Heffley, my associate advisors, for their time, encouragement and expertise.

Other faculty members have also provided helpful scholarly support and encouragement, especially Dr. Thomas J. Miceli, Dr. Francis W. Ahking, Dr. Stephen M. Miller, and Dr. Fred V. Carstensen.

I also want to thank Dr. George M. Katsimbris for his valuable advice to continue my graduate studies at the University of Connecticut.

During my graduate studies at the University of Connecticut, I was extremely fortunate to find individuals, who I consider to be lifelong friends, Georgios Chortareas, Srabani Guha, Luoana Dulgheru, and Jieling Chen.

I thank my sister, Despina Desli, for her love and support. Finally, I thank my parents, Filoktimon Deslis and Athina Desli, and for their love, encouragement, support and guidance in every step of my life. They gave me much and helped me to achieve my dreams, while they asked for little in return. I owe to them who or what I am today. It is to my parents that this study is dedicated.

TABLE OF CONTENTS

List of Tables

CHAPTER 1: INTRODUCTION

1.1 Motivation

There is considerable interest in the relative performance of different production units within an industry, which is measured by the technical efficiency index. The Debreu (1951)-Farrell (1957) measure of technical efficiency is the radial distance between an observed input-output bundle and the frontier of the production possibility set. A firm is called technically efficient when the observed bundle lies on the production frontier. Output-oriented technical efficiency is defined as the gap between the output level obtained by a firm and the maximum level of output which can be produced from a given input bundle. Similarly an input-oriented measure of technical efficiency is defined as the gap between the actual input quantities used and the minimal input quantities required for the production of the observed output.

In order to measure technical efficiency we need to estimate a production function such that all the observed input-output bundles lie on or below the estimated function. In other words, the estimated production function has to be a frontier. Traditional econometric estimation techniques fail to measure a production frontier, because they allow some of the observed output bundles produced by a given set of inputs to be greater than the estimated maximal producible output.

The first attempts to address the problem led to the estimation of deterministic parametric frontier models, non-statistical first (Aigner and Chu, 1968) and statistical later (Afriat, 1972; Richmond, 1974; Greene, 1980). A deterministic frontier treats any deviation from the observed frontier as technical inefficiency and ignores any disturbance due to randomness, specification errors and measurement errors.

Aigner, Lovell and Schmidt (1977) and Meeusen and Van den Broeck (1977) first introduced the concept of a stochastic parametric frontier methodology. Its principal characteristic is a composed error term. The composed error is the sum of a two-sided error term that represents the random shocks and another one-sided error term that represents technical inefficiency. Appropriate methods have been developed for estimation of technical efficiency from these models (Jondrow et al., 1982; Battese and Coelli, 1988). Although this procedure has been extended to panel data, technical efficiency is modeled as an explicit function of time. As a result, one can not distinguish between the technical change and efficiency change. The first essay of the three in the dissertation constructs a model that addresses this problem.

A major drawback of econometric estimation of composed error frontiers is that one is required to make assumptions about the probability distribution of the error terms. Such assumptions of necessity are arbitrary and can lead to different conclusions about the technical efficiency of a firm. The second essay proposes estimation of a parametric stochastic frontier without the need for such assumptions.

A limitation of any parametric frontier (deterministic or statistical) is the subjective choice of the functional form of the frontier. The method of Data Envelopment Analysis (DEA), introduced by Charnes, Cooper and Rhodes (1978, 1981) and extended since then, leads to deterministic non-parametric frontiers. But the resulting technical inefficiency measures are point estimates without any statistical properties. This problem has recently been addressed with the use of bootstrapping. Simar (1992, 1996) and Simar and Wilson

(1997a, 1997b) set the foundation for consistent use of bootstrapping. One problem with this approach is that it assumes that all the firms in the sample have the same probability of getting an observed technical efficiency level. But in reality, the firm's relative efficiency may be systematically influenced by unit specific factors outside the firm's control. The third essay in this dissertation develops a bootstrap procedure that generates the distribution of efficiency for each firm, conditional on unit specific factors.

This chapter provides a brief review of the relevant literature with a focus on the modeling of production frontiers and the measurement of technical efficiency. This review sets the foundation for the development of the essays. Section 1.2 traces the development of the deterministic parametric production frontiers, while section 1.3 deals with the development of stochastic parametric production frontiers. The basic Data Envelopment Analysis (DEA) models are presented in section 1.4. Section 1.5 introduces the basic mechanics of bootstrapping in general. Section 1.6 deals specifically with the existing techniques of the application of the bootstrap to DEA. The last two sections of this chapter explain the main contribution and organization of the rest of this dissertation.

1.2 Deterministic Parametric Frontiers

Consider a firm using k inputs $x=(x_1, x_2,\dots,x_k)$ to produce a single output y. Efficient transformation of the input bundle x into output is characterized by the production function $f(x)$, which shows the maximum producible output from various input vectors. The production function is often characterized as a frontier, because it represents an upper limit of the quantity of output a firm can produce from a given bundle of inputs. For example, assume that the production function has the Cobb-Douglas form and inputs and output bundles are measured in natural logarithms. Then the specified regression model is:

$$y = f(x) + v = \alpha + \sum_{j=1}^{k} \beta_j x_j + v; \quad v \sim N(0, \sigma^2), \tag{1.1}$$

where α is the intercept, $\beta=\{\beta_1, \beta_2, \ldots, \beta_k\}$ are the coefficients of the input vector, and v is the error term that represents the random shocks. Neither the Ordinary Least Squares procedure nor econometric techniques like maximum likelihood estimation restrict the estimated output to be greater than or equal to the observed output levels. Thus, the estimated function is not a production frontier.

1.2.1 Deterministic non-statistical frontiers

Aigner and Chu (1968) develop a mathematical programming method for estimation of a parametric production function with a one-sided error term to ensure that the estimated production function will exceed the output level actually produced from a given bundle of inputs. They specify a Cobb-Douglas production frontier for a sample of n observations:

$$y_i = f(x_i) - u_i = \alpha + \sum_{j=1}^{k} \beta_j x_{ij} - u_i; \quad u_i \geq 0 \quad \text{for } i = 1,2,\ldots,n, \tag{1.2}$$

where the one sided error term u forces all the observed output levels to be on or beneath the frontier ($y \leq f(x)$). The technical inefficiency of each observation is equal to the exponential of the deviation from the frontier u (e^u). The parameters α and $\beta=\{\beta_1, \beta_2, \ldots, \beta_k\}$ are estimated by minimizing the sum of absolute values of the residuals, subject to the constraint that each residual should be non-positive (linear programming):

$$\min \sum_{i=1}^{n} u_i$$

$$\text{st} \quad \alpha + \sum_{j=1}^{k} \beta_j x_{ij} - u_i = y_i \; ; \; i = 1,2,\ldots,n \,. \tag{1.3}$$

$$\beta_j \geq 0; \; j = 1,2,\ldots,k$$

$$u_i \geq 0; \; i = 1,2,\ldots,n.$$

Alternatively the parameters are estimated by minimizing the sum of squared residuals, subject to the same constraint that each residual should be non-positive (quadratic programming):

$$\min \sum_{i=1}^{n} u_i^2$$

$$\text{st} \quad \alpha + \sum_{j=1}^{k} \beta_j x_{ij} - u_i = y_i \; ; \; i = 1,2,\ldots,n \tag{1.4}$$

$$\beta_j \geq 0; \; j = 1,2,\ldots,k$$

$$u_i \geq 0; \; i = 1,2,\ldots,n.$$

A problem with the Aigner and Chu approach is that the estimated frontier is deterministic (non-stochastic) and does not allow any stochastic noise to influence the frontier. As a result, any deviation from the frontier is treated as technical inefficiency. Another limitation of this mathematical programming approach, is that the estimated parameters have no statistical properties even though the input-output data set is only a sample from some underlying population. As an ad hoc adjustment for the possibility of statistical noise, Timmer (1971) extends the Aigner and Chu approach by allowing for an arbitrary percentage of observations to cross the frontier.

1.2.2 Deterministic statistical frontiers

The non-statistical model in (1.2) can be made a deterministic statistical frontier by making assumptions about the distribution of the one-sided error term u. Usually, researchers assume that the one-sided error terms are independently and identically distributed (iid), and that they are independent of the input vector x.

Afriat (1972) discusses the problem of estimating a deterministic parametric production frontier using econometric techniques and suggests that the exponential of the one-sided error term u, e^{-u}, should naturally take values between 0 and 1 because $u \geq 0$ and thus $0 \leq e^{-u} \leq 1$. He proposes a two-parameter beta distribution for e^{-u} and suggests that the model should be estimated using the maximum likelihood method.

Richmond (1974) constructs the first econometric model of a frontier, where he specifies the gamma distribution for the disturbance term u. Richmond suggests a method of estimation, which is based on Ordinary Least Squares. The intercept of model (1.2) is adjusted by the mean of the one-sided error terms, $\mu = E(u)$, in order to get a corrected error term that is centered on zero.

$$y_i = (\alpha - \mu) + \sum_{j=1}^{k} \beta_j x_{ij} - (u_i - \mu) = \alpha^* + \sum_{j=1}^{k} \beta_j x_{ij} + u_i^*; \quad \text{where} \quad \mu = E(u) \ \text{for } i = 1,2,\ldots,n \tag{1.5}$$

Note that the corrected error term $u^* = -(u-\mu)$ does not have the normal distribution. The transformed model, which is also called Modified OLS (MOLS), can be easily estimated by OLS to obtain consistent estimates of the parameters α^* and βs. For the gamma distribution, the unbiased estimator of the variance of u^* ($\hat{\sigma}_u^2$) coincides with the similar estimator of the mean. Thus, $\alpha^* + \hat{\sigma}_u^2$ is an estimator of the intercept α. One problem with this approach is that the correction to the intercept depends on the distribution assumed for u and the estimates of the technical efficiency may lead to different conclusions (Førsund, Lovell and Schmidt, 1980). Moreover, the estimated frontier still allows some observed points to be above the fitted production function. Finally, the estimated model provides information only about the *average efficiency* of the industry and not the individual efficiency. However, one can estimate production frontiers for sub-samples of data (e.g. different industries) and then compare their average efficiency levels.

Schmidt (1976) has shown that, if the one-sided error term u follows the exponential or the half-normal distribution, then Aigner and Chu's linear and quadratic programming procedures are equivalent to maximum likelihood estimation. However, as Schmidt points out, the regularity conditions for consistency or asymptotic efficiency of the maximum likelihood estimators are violated for the exponential and half-normal cases.

Greene (1980a) provides us with the conditions that the density of the one-sided error term u must satisfy for the maximum likelihood estimators to have the desirable asymptotic properties. Additionally, he proposes a model based on the gamma density, which satisfies the derived conditions.

Greene (1980b) develops the corrected OLS estimators (COLS) for the deterministic frontier model. He proposes shifting the OLS-estimated intercept by the maximum error obtained from the OLS, $max\ u_i\ \{u_i,\ i=1,2,\ldots,n\}$. The resulting frontier is parallel to the OLS-estimated function and the estimated corrected intercept is consistent. Unlike Richmond's

MOLS frontier, the resulting frontier from COLS provides us with an estimated frontier where all the observed points lie on or below the frontier.

But, a common problem across all the deterministic production frontier methods is that they ignore the statistical disturbance due to randomness, specification errors and measurement errors. For the parametric frontiers, use of a composed error term takes into account this problem.

1.3 Stochastic Parametric Frontiers

Consider a model like model (1.1) where the production function *f(x)* is expressed in logarithmic terms and it provides us with the efficient transformation of the input vector x into the maximal producible output y^f.

$$y^f_i = f(x_i;\beta) + v_i;\ \ v_i \sim N(0,\sigma_v^2)\ \text{ for } i = 1,2,\ldots,n, \tag{1.6}$$

where v_i represents random shocks to the frontier. If the i-th firm was not be able to produce y^f_i, but instead the observed $y_i \leq y^f_i$, then the deviation from the frontier is due to technical inefficiency. Let $u_i = y^f_i - y_i \geq 0$ denote the deviation from the frontier. Now the model in (1.6) can be written as a stochastic production frontier:

$$\begin{aligned} y_i &= y^f_i - u_i = f(x_i;\beta) + v_i - u_i; \\ & v_i \sim \text{iid } N(0,\sigma_v^2) \text{ and } u_i \geq 0\ (u_i \sim \text{iid})\ \text{ for } i = 1,2,\ldots,n, \end{aligned} \tag{1.7}$$

or equivalently

$$\begin{aligned} y_i &= f(x_i;\beta) + \varepsilon_i;\ \ \varepsilon_i = v_i - u_i; \\ & v_i \sim \text{iid } N(0,\sigma_v^2) \text{ and } u_i \geq 0\ (u_i \sim \text{iid})\ \text{ for } i = 1,2,\ldots,n. \end{aligned} \tag{1.8}$$

The error term ε of the above model is the sum of two independent components and is called a composed error term. The first component is a two-sided error term that represents the random shocks and it is identically and independently distributed with a normal probability density function. The second component represents the deviation from the frontier due to technical inefficiency and is one-sided. It is identically and independently distributed with a one-sided density function. The model in (1.8) for stochastic parametric frontiers is usually estimated with the maximum likelihood method and requires the researcher to make assumptions about the distribution of the one-sided error term.

Once the model is estimated the technical inefficiency is the ratio of the exponential of the deviation of the observed output from the expected maximal producible output.

$$\text{Technical inefficiency} = TE_i = e^{y_i - y_i^f} = \frac{e^{y_i}}{e^{y_i^f}} = \frac{e^{y_i}}{e^{f(x_i;\beta)+v_i}} \text{ for } i = 1,2,\ldots,n. \quad (1.9)$$

Aigner, Lovell, and Schmidt (1977) and Meeusen and van den Broek (1977) simultaneously introduced stochastic parametric frontier modeling for cross sectional data. Both papers suggest the use of a composed error term ε_i such that,

$$y_i = \alpha + \sum_{j=1}^{k} \beta_j x_{ij} + \varepsilon_i;\ \varepsilon_i = \mathrm{v_i} - \mathrm{u_i} \text{ and } \mathrm{u_i} \geq 0 \text{ for } \mathrm{i} = 1,2,\ldots,\mathrm{n}, \quad (1.10)$$

where, as explained above, $v_i \sim iid\ N(0, \sigma_v^2)$ represents a two sided error term (noise) and u_i is a one-sided error term which represents technical inefficiency and is independent from the noise.

Aigner, Lovell, and Schmidt (1977) specify for u_i the half-normal distribution, i.e. the normal distribution truncated at zero ($u_i \sim iid\ |N(0, \sigma_v^2)|$), and derive the maximum likelihood estimators. They also consider the exponential distribution as an alternative. ALS interpret the ratio of the standard errors of the error terms, $\lambda=\sigma_u/\sigma_v$ as an indicator of the relative variability of the two sources of random error that firms experience. When either σ_u^2 approaches zero or σ_v^2 becomes extremely large, then λ^2 approaches zero, which implies that the symmetric error dominates the determination of composed error and we should ignore the inefficiency term. Similarly, when either σ_v^2 approaches zero or σ_u^2 becomes extremely large, then λ^2 approaches infinity, which implies that the one-sided error becomes the dominant source of random variation in the model.

Meeusen and van den Broek (1977) present the model in (1.10) for the case where the one-sided term follows the exponential distribution and they also use maximum likelihood estimation.

Model (1.10) with the composed error term can be estimated under alternative assumptions for the one-sided error term with the use of the maximum likelihood method. Stevenson (1980) argues that the zero mean assumptions for the half-normal density of the one-sided error term is restrictive and he proposes the normal distribution with non-zero mean truncated at zero $(u_i \sim N(\mu,\sigma^2);\ u_i \geq 0)$. However, Greene (1993) shows that the log-likelihood is ill behaved for unrestricted μ and that the assumption of non-zero μ considerably inflates the standard error of the other parameters. The gamma distribution for the one-sided error term satisfies the regularity conditions for the asymptotic properties and it was introduced by Greene (1980a), Stevenson (1980) and extended by Greene (1990).

Olson, Schmidt and Waldman (1980) discuss the method of moments approach for the case that the density of the one-sided error term is specified as half normal, Harris (1992) assumes that the truncated normal density, and Greene (1993, 1995) specifies the exponential and gamma densities. A frequent problem with the method of moments is the third central moment might be negative. Olson, Schmidt and Waldman (1980) show that the method of moments is better than the maximum likelihood when λ is less than 3.16, but Coelli (1995) with another Monte Carlo study found that Maximum Likelihood Estimation outperforms the method of moments for large λ.[1]

1.3.1 Measuring technical inefficiency

The maximum likelihood procedure provides estimates for the composed error term, ε_i, but without a decomposition of ε_i into separate estimates of the one-sided error term u_i and the two-sided v_i a measure for the technical inefficiency can not be obtained.

Jondrow, Materov, Lovell, and Schmidt (JMLS, 1982) derive the conditional distribution $u|\varepsilon$ when the one-sided error term, u_i, has the half-normal or exponential density and they suggest the estimation of the mean or mode of $u|\varepsilon$ for a measure of the error term that represents technical inefficiency. They show that for the half normal case, the density of the conditional error term is truncated normal $N(\mu^*, \sigma_*^2)$.

However, later Battese and Coelli (1988) show that when the production frontier is estimated in logarithmic terms, as in (1.10), the appropriate measure of technical inefficiency is not $exp(-E(u|\varepsilon))$ as Jondrow, Materov, Lovell, and Schmidt suggest, but $E(e^{-u} | \varepsilon)$.

[1] As an alternative to more traditional econometric estimation techniques like the maximum likelihood method or the method of moments for the estimation of model (1.10), van den Broeck et al. (1992) and Koop et al. (1992) suggest the Bayesian approach.

Greene (1993) shows that the mean of the conditional distribution, $E(u|\varepsilon)$, enables unbiased but not consistent estimation of u_i because the variance of the estimate doesn't converge to zero.

Horrace and Schmidt (1995, 1996) derive the upper and lower bounds of $u|\varepsilon$ and the exponential $e^{-u} | \varepsilon$ under the assumption that the error term representing technical inefficiency follows the half-normal distribution. Bera and Sharma (1996) and Hjalmaron, Kumbhakar and Hesmati (1996) obtain the confidence intervals for $E(u|\varepsilon)$, and Bera and Sharma (1996) obtain the confidence intervals for $E(e^{-u} | \varepsilon)$.

1.3.2 Panel Data

Pitt and Lee (1981) generalize the stochastic production frontier in (1.8) for panel data:

$$y_{it} = f(x_{it}) + \varepsilon_{it}; \quad \varepsilon_{it} = v_{it} - u_{it}; \mathrm{v}_{\mathrm{it}} \sim \mathrm{N}(0, \sigma_v^2) \text{ and } \mathrm{u}_{\mathrm{it}} \geq 0$$
$$\text{for } i = 1,2,\ldots,n \text{ and } \mathrm{t} = 1,2,\ldots,T \qquad (1.11)$$

They estimate three variations of model (1.11). First, they assume that the one-sided error term representing technical inefficiency is time invariant ($u_{it} = u_i$ for t=1,2,...T). Their second variation allows the one-sided error terms to be identically and independently distributed over time. Finally, the third variation relaxes the assumption of independence and allows the one-sided error terms to be correlated across time but to be identically and independently distributed across firms. Also, no correlation is allowed between the error terms and inputs. From their first model one can derive only the average technical efficiency of the firm over the time-period studied. The second model assumes independence of the one-sided error term across time and it does not provide any information about changes in a firm's inefficiency during the time-period of the sample. Finally, their third model can not provide a measure of the technical inefficiency, since it can not be decomposed from the composed error term.

Schmidt and Sickles (1984) provide a variety of estimates for the model in (1.11) for panel data under the assumption that the one-sided error term u_{it} is time invariant and varies only across firms:

$$\begin{aligned} &y_{it} = f(x_{it}) + \varepsilon_{it};\ \ \varepsilon_{it} = v_{it} - u_{it};\\ &v_{it} \sim N(0, \sigma_v^2) \text{ and } u_{it} = u_i \geq 0 \\ &\text{for } i = 1,2,\ldots,n \text{ and } t = 1,2,\ldots,T. \end{aligned} \tag{1.12}$$

They show how their model can be estimated without any assumptions about the distribution of the error terms *i*) using OLS like the Corrected OLS in Richmond (1974), *ii*) as a fixed effects model (Within estimator), *iii*) as a random effects model with the effects uncorrelated with the regressors (Generalized Least Squares estimator), and *iv*) as a random effects model with the effects correlated with some of the regressors (Hausman-Taylor estimator). Also, they suggest the maximum likelihood estimators given independence between the error terms and distributional assumptions (MLE estimator). Finally they provide a number of tests to compare alternative assumptions and models and to help one select the appropriate estimator.

Battese and Coelli (1988) present a generalization of some of the results presented by Jondrow et al. (1982), under the assumption that panel data on sample firms are available. Their model is similar to the Schmidt and Sickles (1984) model. However, they use a truncated normal distribution with a non-zero mean, as proposed by Stevenson (1980), for the technical inefficiency component. Battese and Coelli argue that for the logarithmic production function, an appropriate measure of technical efficiency for firm *i* is

$$\hat{TE}_i = \frac{E[y_{it} \mid u_i, x_{it},\ t=1,2,\ldots,T]}{E[y_{it} \mid u_i = 0, x_{it},\ t=1,2,\ldots,T]} = E\left[e^{-u_i} \mid \hat{\varepsilon}_i\right] = E\left[e^{-\hat{u}_i}\right] \qquad (1.13)$$

rather than $e^{-E[u_i|\hat{\varepsilon}_{it}]}$, as in Jondrow et al (1982). Further, for panel data the above estimate of technical efficiency is unbiased but inconsistent.

Reifschneider and Stevenson (1991) and Huang and Liu (1994) model the frontier inefficiency error component as a time-invariant function of various casual firm specific factors, $\underline{Z}_i$, and a random component, w_i, to test whether some proportion of the frontier departures can be systematically explained.

$$u_i = g(Z_i) + w_i \quad \text{for } i = 1,2,\ldots n, \qquad (1.14)$$

However, the one-sided error term, w_i, which is the unexplained component of the inefficiency error, is treated as time invariant. They apply their model assuming that the density of w_i follows either the truncated normal or the gamma distribution.

The Schmidt and Sickles (1984), Battese and Coelli (1988), Reifschneider and Stevenson (1991), and Huang and Liu (1994) models for panel data restrict the one-sided error term to be time-invariant. Thus, only the average technical inefficiency across firms over time can be estimated and there is no information how it is changing over time. Subsequent researchers have allowed the technical inefficiency to vary over time, but they all model efficiency as a systematic function of time:

$$y_{it} = f(x_{it}) + \varepsilon_{it}; \; \varepsilon_{it} = v_{it} - u_{it};$$
$$v_{it} \sim N(0, \sigma_v^2) \text{ and } u_{it} = g(t) \geq 0 \quad . \qquad (1.15)$$
$$\text{for } i = 1,2,\ldots,n \text{ and } t = 1,2,\ldots,T.$$

Kumbhakar (1990) and Cornwell, Schmidt and Sickles (1990) were the first to propose time-varying inefficiency. Battese and Coelli (1992, 1995), Lee and Schmidt (1993), Ahn, Lee and Schmidt (1994), Kumbhakar and Hesmati (1993) and Hesmati and Kumbhakar (1994) propose alternative models for panel data with time varying technical efficiency.[2] [3]

1.4 Data Envelopment Analysis (DEA)

The main limitation of any parametric production function (econometric or non-statistical) is that the researcher specifies an explicit, and in some cases, quite restrictive, functional form for the technology (Bauer 1990, Greene 1993). Another disadvantage is that transformation of multiple-inputs to multiple-outputs can be handled only in a dual cost or profit function, but this requires the further assumption of cost or profit optimizing behavior respectively. Data Envelopment Analysis (DEA) provides an alternative framework for the estimation of a deterministic non-parametric frontier involving multiple inputs and outputs and is based only on the hypotheses of monotonicity, convexity, and free disposability of inputs and outputs. The DEA procedure was developed by Charnes, Cooper and Rhodes (1978, 1981) (CCR) with focus on computing relative efficiency of different decision making

---

[2] Ahn, Lee and Schmidt (1994) estimate the Lees and Schmidt model with the method of moments as opposed to maximum likelihood in the Lee and Schmidt article.

[3] Battese and Coelli (1995) offer a generalization of Battese and Coelli (1992) and Reifschneider and Stevenson (1991) by modeling the one-sided error term as function of firm specific characteristics, z_{it}: $u_{it} = z_{it}\delta + w_{it}$, $w_{it} \sim N(0, \sigma^2)$ *for* $w_{it} \geq - z_{it}\delta$.

units (DMUs). The basic CCR model was limited to constant-returns to scale technologies, but was extended later by Banker, Charnes and Cooper (1984) (BCC) to allow variable-returns-to-scale. Here we provide a brief outline of the DEA procedure.

Consider an industry producing a vector of m outputs, $y=(y_1, y_2, \ldots, y_m)$, from a vector of k inputs, $x=(x_1, x_2, \ldots, x_k)$. Let the vectors x^i and y^i represent, respectively, the input and output bundles of the i-th firm or decision making unit (DMU). Suppose that the input-output data are observed for n DMUs. Then the technology set can be completely characterized by the production possibility set:

$$T=\{(x,y): y \text{ can be produced from } x\}. \quad (1.16)$$

The technology is assumed to have the following properties:

P_1. All observed input-output bundles are feasible.

$$(x^i, y^i) \in T; \text{ for every } i=1,2, \ldots, n \quad (1.17)$$

P_2. T exhibits free disposability with respect to inputs.

$$(x^0, y^0) \in T \text{ and } x^1 \geq x^0 \Rightarrow (x^1, y^0) \in T \quad (1.18)$$

P_3. T exhibits free disposability with respect to outputs.

$$(x^0, y^0) \in T \text{ and } y^1 \leq y^0 \Rightarrow (x^0, y^1) \in T \quad (1.19)$$

P_4. T is convex

$$(x^0, y^0) \in T \text{ and } (x^1, y^1)) \in T \Rightarrow (\lambda x^0+(1-\lambda)x^1, \lambda y^0+(1-\lambda)y^1) \in T; \; 0 \leq \lambda \leq 1 \quad (1.20)$$

The free disposal convex hull of the observed input-output bundles is the smallest technology set satisfying assumptions P_1 to P_4 (Varian, 1984). For variable returns to scale (VRS), it can be written as:

$$S = [(x,y) : x \geq \sum_{i=1}^{n} \lambda_i x^i ; y \leq \sum_{i=1}^{n} \lambda_i y^i ; \sum_{i=1}^{n} \lambda_i = 1; \lambda_i \geq 0; \mathrm{i} = 1,2,\ldots,\mathrm{n}] \qquad (1.21)$$

S is an inner approximation of the true technology set. If we assume constant returns to scale (CRS), i.e. all radial expansions as well as (non-negative) contractions of feasible input-output combinations are also feasible, then the technology set is the corresponding free disposal conical hull of the observed points. It can be written as:

$$S^* = [(x,y) : x \geq \sum_{i=1}^{n} \lambda_i x^i ; y \leq \sum_{i=1}^{n} \lambda_i y^i ; \lambda_i \geq 0; \mathrm{i} = 1,2,\ldots,\mathrm{n}] \qquad (1.22)$$

Following Farrell (1957), the input-oriented technical efficiency of the j-th DMU under the assumption of Constant Returns to Scale (CRS) can be obtained as $TE = \min\{\theta : (\theta x^j, y^j) \in S\}$. The input-oriented technical efficiency of the j-th firm under the CRS can be computed by solving the following linear programming (LP) problem:

$$\begin{aligned} & \min \theta_{\mathrm{j}} \\ s.t. \quad & \sum_{i=1}^{n} \lambda_i y_t^i \geq y_t^j ; \text{ for } \mathrm{t} = 1,2,\ldots,\mathrm{m} \\ & \sum_{i=1}^{n} \lambda_i x_s^i \leq \theta_j x_s^j ; \text{ for } \mathrm{s} = 1,2,\ldots,\mathrm{k} \\ & \lambda_i \geq 0 \quad \text{for } \mathrm{i} = 1,2,\ldots,\mathrm{n}. \end{aligned} \qquad (1.23)$$

This is the primal form of the Charnes, Cooper, and Rhodes (CCR, 1978) model. Banker, Charnes and Cooper (BCC, 1984) extend this model to the case of Variable Returns to Scale (VRS) for which we solve problem (1.29) with an additional constraint on the λ_i s. The input-oriented technical efficiency of the j-th firm under the VRS can be computed by solving the following linear programming (LP) problem:

$$\min \theta_j$$
$$\begin{aligned} s.t.\ & \sum_{i=1}^{n} \lambda_i y_t^i \geq y_t^j;\ \text{for } t = 1,2,\dots,m \\ & \sum_{i=1}^{n} \lambda_i x_s^i \leq \theta_j x_s^j;\ \text{for } s = 1,2,\dots,k \\ & \sum_{i=1}^{n} \lambda_i = 1; \\ & \lambda_i \geq 0 \quad \text{for } i = 1,2,\dots,n. \end{aligned} \tag{1.24}$$

Similarly, the output-oriented technical efficiency of the j-th firm under CRS and VRS can be computed by solving the linear programming (LP) problems in (1.25) and (1.26) respectively:

$$\max \phi_j$$
$$\begin{aligned} s.t.\ & \sum_{i=1}^{n} \lambda_i y_t^i \geq \phi_j y_t^j;\ \text{for } t = 1,2,\dots,m \\ & \sum_{i=1}^{n} \lambda_i x_s^i \leq x_s^j;\ \text{for } s = 1,2,\dots,k \\ & \lambda_i \geq 0 \quad \text{for } i = 1,2,\dots,n \end{aligned} \tag{1.25}$$

and

$$\max \phi_j$$
$$\begin{aligned} s.t.\ & \sum_{i=1}^{n} \lambda_i y_t^i \geq \phi_j y_t^j;\ \text{for } t = 1,2,\dots,m \\ & \sum_{i=1}^{n} \lambda_i x_s^i \leq x_s^j;\ \text{for } s = 1,2,\dots,k \\ & \sum_{i=1}^{n} \lambda_i = 1; \\ & \lambda_i \geq 0 \quad \text{for } i = 1,2,\dots,n. \end{aligned} \tag{1.26}$$

Following the earliest models, a large number of extensions of the CCR model have been developed, e.g. the free disposal hull model (Deprins, Simar and Tulkens, 1984), cone ratio model (Charnes, Cooper, Huang and Sun, 1990), and the assurance region model (Thompson, Langermeier, Lee, Lee and Thrall, 1990). One of the advantages of the DEA models over the parametric approach is that they can handle multiple-input multiple-output production. However, the main advantage is that no explicit functional form of the production frontier is required. On the downside, its greatest limitation is that the calculated frontier is deterministic and does not accommodate statistical noise in the data.

Several alternative procedures to estimate non-parametric statistical production frontiers are currently available in the literature. DEA was extended to the case of stochastic inputs and outputs through the use of chance–constrained programming (Land, Lovell and Thore, 1988, 1993; Olesen and Petersen 1989, 1995). Banker and Maindiratta (1992) (BM) developed a method to obtain maximum likelihood estimates of a monotone and concave non-parametric production frontier with a composed error term. More recently, Sarath and Maindiratta (1997) proved the consistency of the BM estimators. Banker (1993) showed that in the case of a single-output technology, DEA gives a consistent estimate of the production frontier. He laid the foundation of asymptotic properties of DEA estimators and also constructed an F-test to compare differences in inefficiency distributions across groups of DMUs. As yet another alternative, resampling methods like bootstrap and jackknife have been used to derive the distribution of the DEA efficiency estimators. Recently Simar (1992, 1996), Simar and Wilson (1997a, 1997b) set the foundation for consistent use of bootstrap techniques to generate empirical distributions of efficiency scores and have developed tests of hypotheses relating to returns to scale.

The last essay of this dissertation, which is presented in chapter 4, deals with the development of a consistent bootstrap technique. The next section describes the bootstrap methodology in general. It can be applied in variety of contexts including one in chapter 3 of this dissertation. Section 1.6 includes a summary of the existing bootstrapping techniques as they have been applied in DEA.

1.5 Bootstrap

The idea of the bootstrap was first introduced by Efron (1979), who proposed the use of computer-based simulations to obtain the sampling properties of random variables. The starting point of any bootstrap procedure is a sample of observed data $X=\{x_1, x_2,\ldots,x_n\}$ drawn randomly from some population with an unknown probability distribution f. The basic assumption behind the bootstrap method is that the random sample actually drawn "mimics" its parent population.

1.5.1 Naïve Bootstrap Methodology

Suppose that a sample of observed data $X=\{x_1, x_2,\ldots,x_n\}$ is drawn randomly from some population with an unknown probability distribution f. The sample statistic $\hat{\theta}=\theta(X)$ computed from this state of observed values is merely an estimate of the corresponding population parameter $\theta=\theta(f)$. When it is not possible to analytically derive the sampling distribution of that statistic, one examines its empirical density function. Unfortunately, however, the researcher has access to only one sample rather than multiple samples drawn from the same population. As noted above the basic assumption behind the bootstrap method is that the random sample actually drawn "mimics" its parent population. Therefore, if one draws a random sample with replacement from the observed values in the original sample, it can be treated like a sample drawn from the underlying population itself. Repeated samples with replacement yield different values of the sample statistic under investigation and the associated empirical distribution (over these samples) can provide the sampling distribution of this statistic. For reasons explained later this is known as a naïve bootstrap.

The bootstrap sample $X^*=\{x_1^*, x_2^*,\ldots,x_n^*\}$ is an unordered collection of n items drawn randomly from the original sample X with replacement, so that any x_i^* $(i=1,2,\ldots,n)$ has $1/n$ probability of being equal to any x_j $(j=1,2,\ldots,n)$. Some observations from the original sample X will appear zero times in the bootstrap sample, while other observations will appear more than one time. Let $\hat{f}$ denote the empirical density function of the observed sample X from which X^* was drawn. Then it can take the form:

$$\hat{f}(t)=\begin{cases}1/n & \text{if } t=x_i^*, i=1,2,\ldots,n \\ 0 & \text{otherwise}\end{cases}. \qquad (1.27)$$

If $\hat{f}$ is a consistent estimator of f, then the bootstrap distributions will mimic the original unknown sampling distributions of the estimators that we are interested in. Let $\hat{\theta}^* = \theta(X^*)$ be the estimated parameter from the bootstrap sample X^*. Then the distribution of $\hat{\theta}^*$ around $\hat{\theta}$ in $\hat{f}$ is the same as of $\hat{\theta}$ around θ in f. That is:

$$(\hat{\theta}^* - \hat{\theta}) \mid \hat{f} \sim (\hat{\theta} - \theta) \mid \hat{f}. \qquad (1.28)$$

Since every time that we replicate the bootstrap sample we get a different sample X^*, we will also get a different estimate of $\hat{\theta}^* = \theta(X^*)$. Thus we need to select a large number of bootstrap samples, B, in order to extract as many combinations of x_j $(j=1,2,\ldots,n)$ as possible. The bootstrap algorithm has the following steps:

i) Compute the statistic $\hat{\theta} = \theta(X)$ from the observed sample X.

ii) Select b-th ($b=1,2,\ldots,B$) independent bootstrap sample X_b^*, which consists of n data values drawn with replacement from the observed sample X.

iii) Compute the statistic $\hat{\theta}^* = \theta(X_b^*)$ from the b-th bootstrap sample X_b^*.

iv) Repeat steps (ii)-(iii) a large number of times (say B times).

v) Calculate the average of the bootstrap estimates of θ as the arithmetic mean

$$\hat{\theta}^*(\cdot) = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b^* . \qquad (1.29)$$

A measure of the accuracy for the estimator $\hat{\theta}$ as an estimate of θ is the bias, which is defined as the difference between the expectation of $\hat{\theta}$ and θ.

$$\mathrm{bias}_f = \mathrm{bias}_f(\hat{\theta}, \theta) = E_f(\hat{\theta}) - \theta . \qquad (1.30)$$

If the bias is a positive number, then the estimator overestimates the true parameter. If the bias is a negative number, then the estimator underestimates the true parameter. An unbiased estimator will have zero bias, i.e. $E_f(\hat{\theta}) = \theta$. One can use the estimate of bias to bias-correct the estimator, so it becomes less biased. The bias-corrected estimator is

$$\hat{\theta}_{bc} = \hat{\theta} - \mathrm{bias}_f . \qquad (1.31)$$

Similarly, we can estimate the bias of the bootstrap estimator $\hat{\theta}^*_b$, (b = 1,2,..., B) as an estimate of $\hat{\theta}$, $\text{bias}_{\hat{F}} = E_{\hat{F}}(\hat{\theta}^*_b) - \hat{\theta}$. We approximate the expectation of each bootstrap estimator $\hat{\theta}^*_b$ by the average of the bootstrap estimators $\hat{\theta}^*(\cdot)$. Now the estimated bias of the bootstrap estimator based on *B* replications is

$$\text{bias}_B = \hat{\theta}^*(\cdot) - \hat{\theta}. \tag{1.32}$$

Taking bias_B as an estimate for the unknown bias_F, the bias-corrected estimator of θ is

$$\hat{\theta}_{bc} = \hat{\theta} - \text{bias}_B = 2\hat{\theta} - \hat{\theta}^*(\cdot). \tag{1.33}$$

Notice that if $\hat{\theta}^*(\cdot)$ is greater than $\hat{\theta}$, then the bias-corrected estimate $\hat{\theta}_{bc}$ should be less than $\hat{\theta}$. This is the case because we are using the relationship between $\hat{\theta}^*(\cdot)$ and $\hat{\theta}$ to mimic the relationship between $\hat{\theta}$ and θ.

Efron and Tibshirani (1993) point out that bias correction can be problematic in some situations. Even if $\hat{\theta}^*_{bc}$ is less biased than $\hat{\theta}$, it might have substantial by greater standard error due to high variability in bias_B. The standard error of $\hat{\theta}^*(\cdot)$ is measured as

$$\text{se}_B = \text{se}(\hat{\theta}^*) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}^*_b - \hat{\theta}^*(\cdot))^2}\ . \tag{1.34}$$

Correcting for the bias may result in a larger root mean squared error. If bias_B is small compared to the estimated standard error of $\hat{\theta}^*(\cdot)$, then it is safer to use $\hat{\theta}$ than $\hat{\theta}_{bc}$. As a rule of thumb Efron and Tibshirani (1993) suggest the computation of the ratio of the estimated bootstrap bias to standard error, $\text{bias}_B/\text{se}_B$. If the bias is less than .25 standard errors, then it can be ignored.

Finally, we can obtain the bias-corrected estimator from each bootstrap $\hat{\theta}^*_{b,bc}$, (b = 1,2,..., B). We want the corrected empirical density function of $\hat{\theta}^*_b$, (b = 1,2,..., B) to be centered on $\hat{\theta}_{bc}$, the bias-corrected estimate of θ, i.e. $E(\hat{\theta}^*_{b,bc}) = \hat{\theta}_{bc}$, (b = 1,2,..., B). According to this, the bias-corrected estimate from each bootstrap will be

$$\hat{\theta}^*_{b,bc} = \hat{\theta}^*_b - 2\,\text{bias}_B,\ (b = 1,2,\ldots, B). \tag{1.35}$$

If we have corrected by only 1*bias_B, then we would have centered the empirical distribution of $\hat{\theta}^*_b$, (b = 1,2,..., B) on $\hat{\theta}$ instead of $\hat{\theta}_{bc}$. The correction by 1*bias_B will be appropriate for the case where the bias_B is small compared to the estimated standard error of $\hat{\theta}^*(\cdot)$, i.e. the bias-corrected estimate from each bootstrap should be

$$\hat{\theta}^*_{b,bc} = \hat{\theta}^*_b - \text{bias}_B,\ (b = 1,2,\ldots, B). \tag{1.36}$$

Once we have the bias-corrected estimates we can use the percentile method to construct the (1-2a)% confidence intervals for θ as

$$(\hat{\theta}^{*(a)}_{bc}, \hat{\theta}^{*(1-a)}_{bc}),\ (b = 1,2,\ldots, B), \tag{1.37}$$

where $\hat{\theta}^{*(a)}_{bc}$ is the (100*a^{th}) percentile of the empirical density of $\hat{\theta}^*_{b,bc}$, (b = 1,2,..., B).

One major drawback of the bootstrap procedure outlined is that even when sampling with replacement, a bootstrap sample will not include observations from the parent population that were not drawn in the initial sample. The shape of the empirical distribution $\hat{f}$ has jumps at the observed points and it looks like a collection of boxes of width *h*, a small number, centered at the observations and zero anywhere else. Thus, the bootstrap samples are effectively drawn from a discrete population and they fail to reflect the fact that the underlying population density function *f* is continuous. Hence, the empirical distribution from the bootstrap samples as they were drawn in this section is an inconsistent estimator of the population density function. This is why it is known as a naïve bootstrap.

1.5.2 Smooth Bootstrap methodology

One way to overcome this problem is to use kernel estimators as weight functions. The empirical distribution $\hat{f}$ will take the form:

$$\hat{f}(t) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{t - x_i}{h}\right), \tag{1.38}$$

where *h* is the window width or smoothing parameters for the density function. *K(.)* is a kernel function, which satisfies the condition

$$\int_{-\infty}^{\infty} K(x)dx = 1. \tag{1.39}$$

Usually K is a symmetric probability density function like the normal density function. If we use the standard normal density function as the Kernel density function, then the smoothing is called Gaussian smoothing. The empirical density function then can be written as

$$\hat{f}(t) = \frac{1}{nh}\sum_{i=1}^{n}\phi\left(\frac{t - x_i}{h}\right). \tag{1.40}$$

Here $\phi(.)$ is the standard density function.

By virtue of the convolution theorem (Efron and Tibshirani, 1993) we can generate the smoothed bootstrap sample $X^{**}=\{x_1^{**}, x_2^{**}, ..., x_n^{**}\}$ as

$$x_i^{**} = x_i^{*} + h\,\varepsilon_{i,v} \sim f;\ \ i=1,2,...,n, \tag{1.41}$$

where x_i^* is from the naïve bootstrap sample in the previous section.

Sometimes it is the case that the natural domain of the definition of the density function to be estimated is not the whole real line but an interval bounded on one side or both sides. For example we might be interested in obtaining density estimates $\hat{f}$ for which $\hat{f}(x)$ is zero for all negative x. One possible way to solve the above problem is to calculate $\hat{f}(x)$ ignoring the boundary restrictions and then to set the empirical density function equal to zero for values of x that are out of the boundary domain. A drawback of this approach is that the estimates of the empirical density function will no longer integrate to unity.

Silverman (1986) suggests the use of the negative reflection technique to handle such problems. Suppose that we are interested in values of x such that $x \geq \alpha$. If the resulting value from the bootstrap is $x_i^{**} < \alpha$, then we will reflect the x_i^{**}, such that $2\alpha - x_i^{**} \geq \alpha$. The empirical density function will be:

$$\hat{f}(t) = \frac{1}{nh}\sum_{i=1}^{n}\left[\phi\left(\frac{t - x_i}{h}\right) + \phi\left(\frac{t - 2\alpha + x_i}{h}\right)\right]. \tag{1.42}$$

Again by the convolution theorem we can generate the smoothed bootstrap sample $X^{**}=\{x_1^{**}, x_2^{**},\ldots,x_n^{**}\}$ as

$$x_i^{**} = \begin{cases} x_i^* + h\varepsilon_i & \sim \dfrac{1}{nh}\sum_{i=1}^{n}\phi\left(\dfrac{t-x_i}{h}\right) & \text{if } x_i^* + h\varepsilon_i \geq \alpha \\ 2\alpha - (x_i^* + h\varepsilon_i) & \sim \dfrac{1}{nh}\sum_{i=1}^{n}\phi\left(\dfrac{t-2\alpha+x_i}{h}\right) & \text{otherwise} \end{cases} \tag{1.43}$$

where x_i^* is from the naïve bootstrap sample in the previous section.

Choice of the smoothing parameter (h) is crucial to the estimated empirical density function. Following Silverman (1986) we can select the value of the window width that minimizes the approximate mean integrated square error. This leads to

$$h = 0.9\, A\, n^{-1/5}, \tag{1.44}$$

where $A = min$ (standard deviation of X, interquartile range of $X/1.34$),

The bootstrap algorithm can be re-written as following:

i) Compute the statistic $\hat{\theta} = \theta(X)$ from the observed sample X.

ii) Select b-th ($b=1,2,\ldots,B$) independent naive bootstrap sample $X_b^*=\{x_{1,b}^*, x_{2,b}^*,\ldots,x_{n,b}^*\}$, which consists of n data values drawn with replacement from the observed sample X.

iii) Construct the smoothed bootstrap sample $X_b^{**}=\{x_{1,b}^{**}, x_{2,b}^{**},\ldots,x_{n,b}^{**}\}$, from the naïve bootstrap sample according to (1.47) or (1.49).

iv) Compute the statistic $\hat{\theta}^* = \theta(X_b^*)$ from the b-th bootstrap sample X_b^*.

v) Repeat steps (ii)-(iii) a large number of times (say B times).

vi) Calculate the average of the bootstrap estimates of θ as the arithmetic mean

$$\hat{\theta}^*(\cdot) = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}_b^* . \tag{1.45}$$

We can now calculate the bias, bias-corrected estimates and construct confidence intervals following the same steps described in section 1.5.1.

1.6 DEA and Bootstrap

Recently Simar (1992, 1996), Simar and Wilson (1997a, 1997b) set the foundation for the consistent use of bootstrap techniques to generate empirical distributions of efficiency scores and have developed tests of hypotheses relating to returns to scale of bootstrapping. Following Simar and Wilson (1997a) we can describe the existing bootstrap techniques for the output-oriented technical efficiency measure given in (1.32) with the following algorithm:

i) Solve (1.32) to obtain $\hat{\phi}_j$ for each DMU j=1,2,...,n.

ii) Select the b-th (b=1,2,...,B) independent naive bootstrap sample $\{\phi_{1.b}^*, \phi_{2.b}^*, ..., \phi_{n.b}^*\}$, which consists of n data values drawn with replacement from the estimated values $\hat{\phi}_j$ s.

iii) Construct the smoothed bootstrap sample $\{\phi_{1.b}^{**}, \phi_{2.b}^{**}, ..., \phi_{nb}^{**}\}$, from the naïve bootstrap sample. Notice that all the ϕ_j s are greater than or equal to 1. Therefore, the smoothed bootstrap sample should be appropriately bounded. It will be computed according to:

$$\phi_{j,b}^{**} = \begin{cases} \phi_j^{*} + h\varepsilon_j & \text{if } \phi_j^{*} + h\varepsilon_j \geq 1 \\ 2-(\phi_j^{*} + h\varepsilon_j) & \text{otherwise} \end{cases} ;\ \text{for } j = 1,2,\ldots,n. \tag{1.46}$$

As before, h is the optimal width that minimizes the approximate mean integrated square error of $\hat{\phi}_j$ s distribution.

$$h = 0.9\,A\,n^{-1/5},$$

where $A = min$ (standard deviation of ϕ, interquartile range of $\phi/1.34$) (1.47)

iv) Create the b-th pseudo-data set as $\{(x^{j*}, y_j^{*} = y_j\, \hat{\phi}_j / \phi_j^{**}\ );\ j=1,2,\ldots,n\}$.

v) Use the pseudo-data set to compute new $\hat{\varphi}_j^{*}$s from the linear program described in (1.32).

vi) Repeat steps (ii)-(iv) B-times to obtain $\{\hat{\varphi}_{j,b}^{*};\ b=1,2,\ldots,B\}$ for each DMU j, $j=1,2,\ldots,n$.

vii) Calculate the average of the bootstrap estimates of ϕ s, the bias and the confidence intervals as they are described in the previous section.

A problem with the existing bootstrap techniques is that they assume that all DMUs have the same probability of drawing a specific efficiency score. This may not be true when unit-specific factors systematically influence the efficiency level of the DMU but cannot be included in the DEA model. The third essay in this dissertation develops a bootstrap procedure that generates the distribution of efficiency for each individual DMU conditional on unit specific factors.

1.7 Main contributions of the dissertation

As discussed in the previous sections, traditional econometric estimation techniques fail to measure a production frontier, because they allow the observed output bundle produced by a given set of inputs to be greater than the estimated maximal producible output. The first and the second essay deal with the estimation of parametric production frontiers while the third essay deals with a non-parametric production frontier. While the first essay uses econometric techniques to estimate a statistical frontier from the beginning, the last two essays first estimate non-statistical production frontiers and then develop a framework to obtain statistical estimates. Each essay contributes to the literature on estimation of statistical production frontiers and technical efficiency using alternative techniques.

The first essay of the three in the dissertation constructs a model specifying the efficiency change through firm-specific intercepts that evolve over time as a first order auto-regressive process. Apart from allowing efficiency in one period to be influenced by past levels of efficiency, this approach separates efficiency from technical change.

The second essay returns to the Aigner and Chu (1968) approach of modeling a deterministic frontier using mathematical programming, but estimates a parametric production function with a composed error term instead of a one-sided error term. The individual levels of technical efficiency can be estimated without any distributional assumptions about the statistical distribution of the error terms. But this mathematical programming approach yields only point estimates for the parameters of interest. In order to overcome this problem, bootstrapping techniques are employed and confidence intervals for our parameters are constructed.

A criticism for any parametric frontier estimation is the subjective choice of the functional form of the frontier. The proposed alternative, Data Envelopment Analysis, provides point estimates of the relative technical efficiency of a firm. Bootstrapping has been used to generate the distribution of the technical efficiency. A problem with the existing approach is that it assumes that all the firms in the sample have the same probability to get an observed technical efficiency level, while the firm's relative efficiency position might be systematically influenced by unit specific factors out of the firms control. The third essay in this dissertation develops a bootstrap procedure that generates the distribution of efficiency for each firm, conditional on unit specific factors.

1.8 Organization of the dissertation

The main body of this dissertation is contained in Chapters 2 to 4. The first essay, Chapter 2, relates to the econometric estimation of production frontiers for panel data and modeling of technical efficiency and technical change. The second essay, Chapter 3, develops a methodology for the estimation of a statistical production frontier using mathematical programming and bootstrapping techniques. The last essay, Chapter 4, estimates a non-parametric frontier using Data Envelopment Analysis and bootstrapping. Each essay includes a methodological extension and an empirical application. Finally, Chapter 5 contains a summary of the main findings of the dissertation.

CHAPTER 2: FRONTIER PRODUCTION FUNCTION MODELS WITH AUTOREGRESSIVELY TIME-VARYING EFFICIENCY

2.1 Introduction

Econometric estimation of production frontiers started with deterministic production frontiers (Richmond, 1974; Greene, 1980) and was later extended to stochastic production frontiers in a pair of seminal papers by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). The basic model for the estimation of a stochastic parametric production function includes a composed error term, which is the sum of an exogenous shock represented by a two-sided error term and technical efficiency represented by a one-sided error term. Subsequently, Jondrow, Materov, Lovell and Schmidt (1982) and Battese and Coelli (1988) showed how individual efficiency levels can be measured from such a model.

At the initial stage of development of the stochastic production frontier, the specified models were limited to cross sectional data. Pitt and Lee (1981), Schmidt and Sickles (1984), Reifschneider and Stevenson (1991) and Huang and Liu (1994) provided applications using panel data. However, their models treated technical efficiency as time invariant. Subsequent researchers have allowed the technical efficiency of a firm to vary over time, but they model efficiency as a systematic function of time (Kumbhakar 1990; Cornell, Schmidt and Sickles, 1990; Battese and Coelli, 1992, 1995; Lee and Schmidt, 1993; Kumbhakar and Hesmati, 1993; Hesmati and Kumbhakar, 1994). The problem with this approach is that, in most econometric models using time series data, technical change is also specified as an explicit function of time. As a result, one can not distinguish between technical change and efficiency change in these models.

This chapter constructs a model specifying efficiency change through firm-specific intercepts that evolve over time as a first order auto-regressive process (AR(1)). This is consistent with the belief that people learn from mistakes gradually. This approach builds on the Cooley-Prescott (1973a, 1973b) adaptive regression model within the class of regression models with time-varying parameters. Apart from allowing efficiency in one period to be influenced by past levels of efficiency, the developed model separates technical efficiency from technical change.

This essay is organized as follows. The adaptive regression model is presented in section 2.2.1 and in section 2.2.2 we set up our specific model. Section 2.3 contains several sub-sections that explain how the model will be estimated, define the likelihood function, derive the first order conditions and the information matrix for a maximum likelihood function, and finally describe how technical efficiency at individual data points can be measured from the estimated model. An empirical application is provided in section 2.4. Section 2.5 contains a summary of the contributions of the second chapter.

2.2 Frontier Production Function Model with Autoregressively Time-Varying Efficiency

In this paper we take an alternative approach to model technical efficiency for panel data and conceptualize change in technical efficiency of a firm as a first order autoregressive

(AR(1)) process. This builds on the Cooley-Prescott adaptive regression model, which is one form of regression model with time-varying parameters.

2.2.1. Adaptive Regression Models

The Cooley-Prescott adaptive regression model can be specified as:

$$\begin{aligned} y_t &= \alpha_t + \beta x_t + v_t; \quad v \sim \text{iid } N(0,\sigma_v^2), \\ \alpha_t &= \alpha_{t-1} + u_t; \quad u \sim \text{iid } N(0,\sigma_u^2) \quad \text{for } t = 1,2,\ldots,T. \end{aligned} \tag{2.1}$$

The error terms v_t and u_s in the above model are assumed to be independently distributed for all t and s (Cooley and Prescott, 1973). The term adaptive regression is due to the fact that the residual in (2.1) is the sum of a random walk, α_t, and an independent error, v_t.[1] The major advantage of the Cooley and Prescott model is that the change in the constant term takes into account structural changes and thereby obtains better forecasts.[2]

Rosenberg (1973) considered a model similar to that of Cooley and Prescott. Instead of making the coefficients β a random walk, he specified a stochastically convergent parameter structure. The Rosenberg model applied on the intercept changes becomes:

$$\begin{aligned} y_t &= \alpha_t + \beta x_t + v_t; \quad v \sim \text{iid } N(0,\sigma_v^2), \\ \alpha_t &= (1-\lambda)\alpha + \lambda\alpha_{t-1} + u_t; \quad u \sim \text{iid } N(0,\sigma_u^2) \quad \text{for } t = 1,2,\ldots,T. \end{aligned} \tag{2.2}$$

---

[1] They also allow the coefficient parameter β to vary in a similar manner like the intercept in (2.1) and they call this model the varying-parameter regression model.

[2] Maddala (1982) argues that the adaptive regression model captures the effect of omitted variables.

Both models can be estimated by generalized least squares or maximum likelihood procedures after deriving the covariance matrix of the residual.

2.2.2. The Model with Autoregressively Varying Efficiency

Consider a panel data set from n firms observed over T periods. Let y_{it} and x_{it} represent respectively the scalar output level and the input vector of k inputs for firm i at time t. The production function is specified as:

$$y_{it} = \alpha_{it} + x_{it}\beta + v_{it}, \qquad (2.3)$$

where v_{it} is the error term that represents random shocks, β is the vector of k parameters for the input vector. Finally, α_{it} is the firm specific intercept that evolves over time as an autoregressive (AR(1)) process:

$$\alpha_{it} = \alpha_i + \phi\alpha_{i,t-1} + u_{it}; \quad u_{it} \leq 0. \qquad (2.4)$$

u_{it} is the error term that it is due to technical inefficiency.

Since the technical inefficiency is introduced into the model through the intercept and not as a deterministic function of time, we can include time as one of the explanatory variables in the vector x_{it}. This allows us to distinguish between technical change and efficiency change.

If the i-th firm is 100% efficient at time t, then u_{it} will be equal to zero. The firm's intercept at time t, α_{it}, depends on last period's intercept and takes the maximum value:

$$\alpha_{it}^{100} = \alpha_i + \phi\alpha_{i,t-1}. \qquad (2.5)$$

Notice that the firm need not have been fully efficienct at time t and its intercept for the previous period does not have to be at its maximum level. *If the firm is 100% efficient at all* *points in time*, then all the error terms u_{it} for firm *i* will be zero and the intercept at any point of time will evolve as:

$$\alpha_{it}^{100} = \alpha_i + \phi\alpha_{i,t-1}^{100}. \tag{2.6}$$

The maximum long run intercept for the efficient firm will be:

$$E(\alpha_{it}^{100}) = \frac{\alpha_i}{1-\phi}. \tag{2.7}$$

If the i-th firm is inefficient at time *t*, then u_{it} is less than zero and the firm's intercept at time *t*, α_{it}, is going to be less than its maximum value. The long run intercept for firm *i* will be:

$$E(\alpha_{it}) = \frac{\alpha_i + \mu}{1-\phi}; \quad \mu = E(u_{it}) < 0. \tag{2.8}$$

As shown in equation (1.9), the technical inefficiency of firm *i* at time *t* is the ratio of the exponential of the deviation of the observed output, y_{it}, from the expected maximal producible output, y_{it}^f, i.e.:

$$TE_{it} = e^{y_{it} - y_{it}^f} = e^{u_{it}}. \tag{2.9}$$

We can introduce into the model other variables, z_{it} that can influence the position of the intercept by specifying

$$\alpha_{it} = \alpha_i + \phi\alpha_{i,t-1} + z_{it}\gamma + u_{it}; \quad u_{it} \le 0, \tag{2.10}$$

where γ is the vector of coefficients that correspond to variables z_{it}. The long run intercept for firm firm i will be:

$$E(\alpha_{it}) = \frac{\alpha_i + z_{it}\gamma + \mu}{1-\phi}; \quad \mu = E(u_{it}) < 0.$$

For our application we specify time as the explanatory variable z_{it}, in order to capture technical change. This feature allows both technical change and technical efficiency to influence the position of the intercept, while we obtain measures for both.

Returning to the model described by equations (2.3) and (2.10), we can rewrite the model in the following simplified form:

$$y_{it} = \alpha_i + \phi y_{i,t-1} + x_{it}\beta - \phi x_{i,t-1}\beta + z_{it}\gamma + (v_{it} - \phi v_{it-1} + u_{it}). \tag{2.11}$$

or

$$\begin{gathered} y_{it} = \alpha_i + \phi y_{i,t-1} + x_{it}\beta - \phi x_{i,t-1}\beta + z_{it}\gamma + \varepsilon_{it} \\ \text{with } \varepsilon_{it} = (v_{it} - \phi v_{it-1} + u_{it}) \end{gathered}. \tag{2.12}$$

The composed error term, ε_{it}, of this model is the sum of an MA(1) process and a one-sided error term.

2.3 Estimation of the model

Consider a random variable $\tilde{u}_{it} \sim N(0, \sigma_{\tilde{u}}^2)$ and $\tilde{u}_i^* = \max_t\{\tilde{u}_{it}\}$ is the maximum value of a sample $(\tilde{u}_{i1}, \tilde{u}_{i2}, ..., \tilde{u}_{iT})$ for firm i drawn from this distribution. Then we can define:

$$u_{it} = \tilde{u}_{it} - \tilde{u}_i^* = \tilde{u}_{it} - \max_t\{\tilde{u}_{it}\} \leq 0. \quad (2.13a)$$

Thus,

$$\alpha_{it} = \tilde{\alpha}_i + \phi\alpha_{i,t-1} + z_{it}\gamma + \tilde{u}_{it}, \; \tilde{u}_{it} \sim N(0, \sigma_u^2) \quad (2.13b)$$

Now the model with equations (2.3) and (2.13) can be written as:

$$y_{it} = \tilde{\alpha}_i + \varphi y_{i,t-1} + x_{it}\beta - \phi\, x_{i,t-1}\beta + z_{it}\gamma + (v_{it} - \phi v_{it-1} + \tilde{u}_{it}). \quad (2.14)$$

or

$$\begin{aligned} y_{it} &= \tilde{\alpha}_i + \phi y_{i,t-1} + x_{it}\beta - \phi\, x_{i,t-1}\beta + z_{it}\gamma + \tilde{\varepsilon}_{it} \\ &\text{where } \tilde{\varepsilon}_{it} = (v_{it} - \phi v_{it-1} + \tilde{u}_{it}) \end{aligned} \quad (2.15)$$

After we obtain consistent estimates for the parameters of the above model, we can apply the Greene correction method to calculate the corrected intercept and the one-sided error terms (Greene, 1980). The Greene correction will result in the estimation of a consistent but biased estimate of the frontier intercept and the corrected intercept, and the one-sided residuals will be given by:

$$\begin{aligned} \hat{\alpha}_i &= \hat{\tilde{\alpha}}_i + \tilde{u}_i^* = \hat{\tilde{\alpha}}_i + \max_t\{\hat{\tilde{u}}_{it}\} \text{ and} \\ \hat{u}_{it} &= \hat{\tilde{u}}_{it} - \tilde{u}_i^* = \hat{\tilde{u}}_{it} - \max_t\{\hat{\tilde{u}}_{it}\} \leq 0 \end{aligned} \quad (2.16)$$

We can consistently estimate the model in (2.15) using Maximum Likelihood Estimation Method, as it is described in the next paragraph. Then we can obtain estimates for the intercept and the one-sided error term using (2.16). The technical efficiency will be given by:

$$TE_{it} = e^{\hat{u}_{it}} = e^{\hat{\tilde{u}}_{it} - \tilde{u}_i^*} = e^{\hat{\tilde{u}}_{it} - \max_t \{\hat{\tilde{u}}_{it}\}} \leq 1 \tag{2.17}$$

2.3.1 Maximum Likelihood Estimation

If we assume that $\tilde{u}_{it} \sim N(0, \sigma_u^2)$ and $v_{it} \sim N(0, \sigma_v^2)$ are mutually uncorrelated orthogonal processes, then $\tilde{\varepsilon}_{it}$ can be written as

$$\tilde{\varepsilon}_{it} = (v_{it} - \phi v_{it-1} + \tilde{u}_{it}). \tag{2.18}$$

Also, $\tilde{\varepsilon}_{it}$ being the sum of an MA(1) and a white noise process would be an MA(1) process.

$$\tilde{\varepsilon}_{it} = w_{it} - \theta w_{i,t-1} \tag{2.19}$$

where $w_{it} \sim N(0, \sigma_w^2)$ is the underlying white noise.

Now the variance and covariance of the composed error term $\tilde{\varepsilon}_{it}$ will be:

$$\begin{aligned} E(\tilde{\varepsilon}_{it}^2) &= (1+\phi^2)\sigma_v^2 + \sigma_u^2 = (1+\theta^2)\sigma_w^2 \\ E(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{it-1}) &= -\phi\sigma_v^2 = -\theta\sigma_w^2 \\ E(\tilde{\varepsilon}_{it}\tilde{\varepsilon}_{it-s}) &= 0 \quad \text{for } s \geq 2 \end{aligned} \tag{2.20}$$

Hence the variance-covariance matrix of the vector with the composed error terms for firm *i*, $\tilde{\varepsilon}_i = \{\tilde{\varepsilon}_{i1}, \tilde{\varepsilon}_{i2}, \ldots, \tilde{\varepsilon}_{iT}\}$ will be:

$$\Sigma = E(\varepsilon_i \varepsilon_i') = \begin{bmatrix} (1+\theta^2)\sigma_w^2 & -\theta\sigma_w^2 & \cdots & 0 \\ -\theta\sigma_w^2 & (1+\theta^2)\sigma_w^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & -\theta\sigma_w^2 \\ 0 & 0 & -\theta\sigma_w^2 & (1+\theta^2)\sigma_w^2 \end{bmatrix} = \sigma_w^2 \Omega \tag{2.21}$$

$$\text{where } \Omega = \begin{bmatrix} 1+\theta^2 & -\theta & \cdots & 0 \\ -\theta & 1+\theta^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & -\theta \\ 0 & 0 & -\theta & 1+\theta^2 \end{bmatrix}.$$

Each vector $\tilde{\varepsilon}_i = \{\tilde{\varepsilon}_{i1}, \tilde{\varepsilon}_{i2}, \ldots, \tilde{\varepsilon}_{iT}\}$ for i=1,2,...,n follows the multivariate normal distribution $\tilde{\varepsilon}_i \sim \text{iid } N(0, \Sigma)$. Thus, the probability density function (p.d.f.) of $\tilde{\varepsilon}_i$ will be

$$f(\tilde{\varepsilon}_i) = \frac{1}{(2\pi)^{T/2}} \frac{1}{|\Sigma|} e^{\frac{1}{2}\tilde{\varepsilon}_i'\Sigma^{-1}\tilde{\varepsilon}_i} \tag{2.22}$$

Since the error vectors $\tilde{\varepsilon}_i$s are independent across firms, the p.d.f. of the vector $(\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n)$ will be

$$f(\tilde{\varepsilon}_1, \tilde{\varepsilon}_2, \ldots, \tilde{\varepsilon}_n) = \frac{1}{(2\pi)^{nT/2}} \frac{1}{|\Sigma|^{n/2}} e^{\frac{1}{2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Sigma^{-1}\tilde{\varepsilon}_i}. \tag{2.23}$$

Now the *likelihood function* can be found by substituting the $\tilde{\varepsilon}_{it}$s with

$$\begin{aligned} \tilde{\varepsilon}_{it} &= y_{it} - \tilde{\alpha}_i + \phi y_{i,t-1} + x_{it}\beta' - \phi\, x_{i,t-1}\beta' \\ &\text{for every } i = 1,2,\ldots,n \text{ and } t = 1,2,\ldots,T. \end{aligned} \tag{2.24}$$

The *Log Likelihood Function* becomes:

$$\begin{aligned} \ln L(\vartheta) &= -\frac{nT}{2}\ln(2\pi) - \frac{nT}{2}\ln(\sigma_w^2)\frac{n}{2}\ln(|\Omega|) - \frac{1}{2\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\tilde{\varepsilon}_i, \\ \text{where } \vartheta &= (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n, \beta_1, \ldots, \beta_k, \gamma_1, \ldots, \gamma_s, \phi, \theta, \sigma_w^2)' \end{aligned} \tag{2.25}$$

where $\tilde{\alpha}=(\tilde{\alpha}_1,\tilde{\alpha}_2,...,\tilde{\alpha}_n)$ is the vector with the firm specific coefficients. The solution of the first order conditions yields estimates of the parameters $\hat{\vartheta}=(\hat{\tilde{\alpha}},\hat{\beta},\hat{\phi},\hat{\theta},\hat{\sigma}_w^2)$. The information matrix permits us to derive the asymptotic standard errors of the estimates.

2.3.2 First Order Conditions

The First Order Conditions can be summarized as:

$$\frac{\partial \ln L}{\partial \tilde{\alpha}_j}=-\frac{1}{\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i{}'\Omega^{-1}\frac{\partial \tilde{\varepsilon}_i}{\partial \tilde{\alpha}_j}=\frac{1}{\sigma_w^2}\tilde{\varepsilon}_j{}'\Omega^{-1}i=0,\ \text{for } j=1,2,\ldots,n, \tag{2.26}$$

where i is a Tx1 vector of ones i=[1,1,...,1]'.

$$\frac{\partial \ln L}{\partial \beta_p}=-\frac{1}{\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i{}'\Omega^{-1}\frac{\partial \tilde{\varepsilon}_i}{\partial \beta_p}=0,\ \text{for } p=1,2,\ldots,k, \tag{2.27}$$

where $\frac{\partial \tilde{\varepsilon}_i}{\partial \beta_p}=-\begin{bmatrix} x_{i1p}-\phi x_{i0p} \\ x_{i2p}-\phi x_{i1p} \\ \vdots \\ x_{iTp}-\phi x_{i,T-1,p} \end{bmatrix}$

$$\frac{\partial \ln L}{\partial \gamma_q}=-\frac{1}{\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i{}'\Omega^{-1}\frac{\partial \tilde{\varepsilon}_i}{\partial \gamma_q}=0,\ \text{for } q=1,2,\ldots,s. \tag{2.28}$$

where $\frac{\partial \tilde{\varepsilon}_i}{\partial \gamma_q}=-\begin{bmatrix} z_{i1q} \\ z_{i2q} \\ \vdots \\ z_{iTq} \end{bmatrix}$

$$\frac{\partial \ln L}{\partial \phi} = -\frac{1}{\sigma_w^2}\sum_{i=1}^{n} \tilde{\varepsilon}_i{}'\Sigma^{-1}\frac{\partial \tilde{\varepsilon}_i}{\partial \phi} = 0, \tag{2.29}$$

where $\frac{\partial \tilde{\varepsilon}_i}{\partial \phi} = -\begin{bmatrix} y_{i0} - x_{i0}\beta \\ y_{i1} - x_{i1}\beta \\ \vdots \\ y_{i,T-1} - x_{i,T-1}\beta \end{bmatrix}$.

$$\frac{\partial \ln L}{\partial \theta} = -\frac{n}{2}\operatorname{trace}(\Omega^{-1}\frac{\partial \Omega}{\partial \theta}) + \frac{1}{2\sigma_w^2}\sum_{i=1}^{n} \tilde{\varepsilon}_i{}'\Omega^{-1}\frac{\partial \Omega}{\partial \theta}\Omega^{-1}\tilde{\varepsilon}_i = 0 \tag{2.30}$$

$$\frac{\partial \ln L}{\partial \sigma_w^2} = -\frac{nT}{2\sigma_w^2} + \frac{1}{2(\sigma_w^2)^2}\sum_{i=1}^{n} \tilde{\varepsilon}_i{}'\Omega^{-1}\tilde{\varepsilon}_i = 0 \tag{2.31}$$

From the last of these First Order Conditions (2.31) we get

$$\sigma_w^2 = \frac{1}{nT}\sum_{i=1}^{n} \tilde{\varepsilon}_i{}'\Omega^{-1}\tilde{\varepsilon}_i , \tag{2.32}$$

which is similar to the weighted sum of Squared Error divided by the number of observations in our sample.

One can use a nonlinear procedure like the Newton-Raphson procedure to solve the system of first order conditions to obtain the Maximum Likelihood Estimators.

2.3.3 Second Order Conditions and the Information Matrix

The second order direct and cross partial derivatives of the likelihood function are:

$$\frac{\partial^2 \ln L}{\partial \alpha_j \partial \alpha_m} = \begin{cases} -\frac{1}{\sigma_w^2} i'\Omega^{-1} i < 0 & \text{if } j = m \\ 0 & \text{if } j \neq m \end{cases}, \quad \text{for } j,m = 1,2,\ldots,n$$

$$\frac{\partial^2 \ln L}{\partial \alpha_j \partial \beta_p} = \frac{1}{\sigma_w^2} i' \Omega^{-1} \frac{\partial \tilde{\varepsilon}_j}{\partial \beta_p}, \quad \text{for } j=1,2,\ldots,n \text{ and } p=1,2,\ldots,k$$

$$\frac{\partial^2 \ln L}{\partial \alpha_j \partial \gamma_q} = \frac{1}{\sigma_w^2} i' \Omega^{-1} \frac{\partial \tilde{\varepsilon}_j}{\partial \gamma_q}, \quad \text{for } j=1,2,\ldots,n \text{ and } q=1,2,\ldots,s$$

$$\frac{\partial^2 \ln L}{\partial \alpha_j \partial \phi} = \frac{1}{\sigma_w^2} i' \Omega^{-1} \frac{\partial \tilde{\varepsilon}_j}{\partial \phi}, \quad \text{for } j=1,2,\ldots,n$$

$$\frac{\partial^2 \ln L}{\partial \alpha_j \partial \theta} = -\frac{1}{\sigma_w^2} \tilde{\varepsilon}_j' \Omega^{-1} \frac{\partial \Omega^{-1}}{\partial \theta} \Omega^{-1} i, \quad \text{for } j=1,2,\ldots,n$$

$$\frac{\partial^2 \ln L}{\partial \alpha_j \partial \sigma_w^2} = -\frac{1}{(\sigma_w^2)^2} \tilde{\varepsilon}_j' \Omega^{-1} i, \quad \text{for } j=1,2,\ldots,n\text{; at the optimal point } \frac{\partial^2 \ln L}{\partial \alpha_j \partial \sigma_w^2} = 0$$

$$\frac{\partial^2 \ln L}{\partial \beta_p \partial \beta_r} = \begin{cases} -\dfrac{1}{\sigma_w^2} \displaystyle\sum_{i=1}^{N} \frac{\partial \tilde{\varepsilon}_i'}{\partial \beta_p} \Omega^{-1} \frac{\partial \tilde{\varepsilon}_i}{\partial \beta_p} \le 0 & \text{if } p = r \\ -\dfrac{1}{\sigma_w^2} \displaystyle\sum_{i=1}^{N} \frac{\partial \tilde{\varepsilon}_i'}{\partial \beta_r} \Omega^{-1} \frac{\partial \tilde{\varepsilon}_i}{\partial \beta_p} & \text{if } p \ne r \end{cases}, \quad \text{for } p,r = 1,2,\ldots,k$$

$$\frac{\partial^2 \ln L}{\partial \beta_p \partial \gamma_q} = -\frac{1}{\sigma_w^2} \sum_{i=1}^{N} \frac{\partial \tilde{\varepsilon}_i'}{\partial \gamma_q} \Omega^{-1} \frac{\partial \tilde{\varepsilon}_i}{\partial \beta_p}, \quad \text{for } p = 1,2,\ldots,k \text{ and } q=1,2,\ldots,s$$

$$\frac{\partial^2 \ln L}{\partial \beta_p \partial \phi} = -\frac{1}{\sigma_w^2} \sum_{i=1}^{N} \frac{\partial \tilde{\varepsilon}_i'}{\partial \phi} \Omega^{-1} \frac{\partial \tilde{\varepsilon}_i}{\partial \beta_p} - \frac{1}{\sigma_w^2} \sum_{i=1}^{N} \tilde{\varepsilon}_i' \Omega^{-1} \frac{\partial^2 \tilde{\varepsilon}_i}{\partial \beta_p \partial \phi}, \quad \text{for } p=1,2,\ldots,k$$

where $\dfrac{\partial^2 \tilde{\varepsilon}_{it}}{\partial \beta_p \partial \phi} = x_{i,t-1,p}$

$$\frac{\partial^2 \ln L}{\partial \beta_p \partial \theta} = \frac{1}{\sigma_w^2} \sum_{i=1}^{N} \tilde{\varepsilon}_i' \Omega^{-1} \frac{\partial \Omega}{\partial \theta} \frac{\partial \tilde{\varepsilon}_i}{\partial \beta_p}, \quad \text{for } p=1,2,\ldots,k$$

$$\frac{\partial^2 \ln L}{\partial\beta_p \partial\sigma_w^2} = \frac{1}{(\sigma_w^2)^2}\sum_{i=1}^{N}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\beta_p}, \quad \text{for p=1,2,...,k; at the optimal point } \frac{\partial^2 \ln L}{\partial\beta_p \partial\sigma_w^2} = 0$$

$$\frac{\partial^2 \ln L}{\partial\gamma_q \partial\gamma_h} = \begin{cases} -\dfrac{1}{\sigma_w^2}\displaystyle\sum_{i=1}^{N}\frac{\partial\tilde{\varepsilon}_i'}{\partial\gamma_q}\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\gamma_q} \le 0 & \text{if } q = h \\ -\dfrac{1}{\sigma_w^2}\displaystyle\sum_{i=1}^{N}\frac{\partial\tilde{\varepsilon}_i'}{\partial\gamma_h}\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\gamma_q} & \text{if } q \ne h \end{cases}, \quad \text{for q,h = 1,2,...,s}$$

$$\frac{\partial^2 \ln L}{\partial\gamma_q \partial\phi} = -\frac{1}{\sigma_w^2}\sum_{i=1}^{N}\frac{\partial\tilde{\varepsilon}_i'}{\partial\phi}\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\gamma_q}, \quad \text{for q=1,2,...,s}$$

$$\frac{\partial^2 \ln L}{\partial\gamma_q \partial\theta} = \frac{1}{\sigma_w^2}\sum_{i=1}^{N}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\Omega}{\partial\theta}\frac{\partial\tilde{\varepsilon}_i}{\partial\gamma_q}, \quad \text{for q=1,2,...,s}$$

$$\frac{\partial^2 \ln L}{\partial\gamma_q \partial\sigma_w^2} = \frac{1}{(\sigma_w^2)^2}\sum_{i=1}^{N}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\gamma_q}, \quad \text{for q=1,2,...,s; at the optimal point } \frac{\partial^2 \ln L}{\partial\gamma_q \partial\sigma_w^2} = 0$$

$$\frac{\partial^2 \ln L}{\partial\phi^2} = -\frac{1}{\sigma_w^2}\sum_{i=1}^{n}\frac{\partial\tilde{\varepsilon}_i'}{\partial\phi}\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\phi}$$

$$\frac{\partial^2 \ln L}{\partial\phi \partial\theta} = \frac{1}{\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\Omega}{\partial\theta}\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\phi}$$

$$\frac{\partial^2 \ln L}{\partial\phi \partial\sigma_w^2} = \frac{1}{(\sigma_w^2)^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\tilde{\varepsilon}_i}{\partial\phi}; \text{ at the optimal point } \frac{\partial^2 \ln L}{\partial\phi \partial\sigma_w^2} = 0$$

$$\frac{\partial^2 \ln L}{\partial\theta^2} = -\frac{n}{2}\frac{\partial^2 \ln|\Omega|}{\partial\theta^2} + \frac{2}{2\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\Omega}{\partial\theta}\frac{\partial\Omega^{-1}}{\partial\theta}\tilde{\varepsilon}_i + \frac{1}{2\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial^2\Omega}{\partial\theta^2}\Omega^{-1}\tilde{\varepsilon}_i =$$

$$= -\frac{n}{2}\text{trace}(-\Omega^{-1}\frac{\partial\Omega}{\partial\theta}\Omega^{-1}\frac{\partial\Omega}{\partial\theta} + 2\Omega^{-1}) - \frac{1}{\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\Omega}{\partial\theta}\Omega^{-1}\frac{\partial\Omega}{\partial\theta}\Omega^{-1}\tilde{\varepsilon}_i + \frac{1}{\sigma_w^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\Omega^{-1}\tilde{\varepsilon}_i$$

$$\frac{\partial^2 \ln L}{\partial\theta\partial\sigma_w^2} = -\frac{1}{2(\sigma_w^2)^2}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\frac{\partial\Omega}{\partial\theta}\Omega^{-1}\tilde{\varepsilon}_i \; ;$$

at the optimal point $\frac{\partial^2 \ln L}{\partial\theta\partial\sigma_w^2} = -\frac{n}{2\sigma_w^2}\,\mathrm{trace}(\Omega^{-1}\frac{\partial\Omega}{\partial\theta})$

$$\frac{\partial^2 \ln L}{\partial(\sigma_w^2)^2} = \frac{nT}{2(\sigma_w^2)^2} - \frac{2}{2(\sigma_w^2)^3}\sum_{i=1}^{n}\tilde{\varepsilon}_i'\Omega^{-1}\tilde{\varepsilon}_i \; ; \text{ at the optimal point } \frac{\partial^2 \ln L}{\partial(\sigma_w^2)^2} = -\frac{nT}{2(\sigma_w^2)^2} \le 0$$

At the optimal point, the resulting Hessian matrix with direct and cross second partial derivatives of the log-likelihood function has to be negative definite for a maximum. These second partials are used to derive the information matrix for the asymptotic variance and covariance standard error of the Maximum Likelihood Estimators $\hat{\vartheta}$:

$$\mathrm{Var}(\vartheta) = -\left[\frac{\partial^2 \ln L}{\partial\vartheta\partial\vartheta'}\right]^{-1}_{\vartheta=\hat{\vartheta}}, \quad \text{where } \vartheta = (\alpha_1,\dots,\alpha_n,\beta_1,\dots,\beta_k,\gamma_1,\dots,\gamma_s,\phi,\theta,\sigma_w^2)'$$

2.3.4 Individual Measures of Technical Efficiency

Based on the previous assumptions, the conditional distribution of $\tilde{u}_{it} \mid \tilde{\varepsilon}_{it}$ is

$\tilde{u}_{it} \mid \tilde{\varepsilon}_{it} \sim N(\mu_{it}^{*}, \sigma_{*}^2)$,

$$\text{where } \mu_{it}^{*} = \frac{\sigma_u^2}{(1+\phi^2)\sigma_v^2 + \sigma_u^2}\tilde{\varepsilon}_{it} \text{ and } \sigma_{*}^2 = \frac{(1+\phi^2)\sigma_v^2\sigma_u^2}{(1+\phi^2)\sigma_v^2 + \sigma_u^2}. \qquad (2.33)$$

Following Battese and Coelli (1988) the technical efficiency will be given by

$$TE_{it} = E(e^{u_{it}} \mid \tilde{\varepsilon}_{it}) = E(e^{\tilde{u}_{it}} \mid \tilde{\varepsilon}_{it})\, e^{-\max_t\{\tilde{u}_{it}\}} = \exp(\mu_{it}^{*} + \frac{\sigma_{*}^{2}}{2} - \max_t\{\tilde{u}_{it}\}). \qquad (2.34)$$

Detailed derivation of the above results is presented in Appendix II.

2.4 Empirical Application

The data used in the empirical application are from the Annual Census of Manufacturing for the period 1953-83, and relate to 12 selected states of India. Table 2.1 lists the states included in our sample. Output is measured by total value added in manufacturing. The two inputs included are labor (L) and capital (K). Labor is measured by persons employed while capital is measured by book value of fixed assets. Nominal valued added has been deflated by the manufactured goods price index. Similarly, book value of fixed capital has been deflated by the price index of machinery and transport equipment. Further, the Census data are state level aggregates. These have been divided by appropriate numbers of establishments covered to measure output and input quantities per establishment in each state. Over this period, reorganization of states has led to several redefinitions of geographical boundaries. As a result, what we have is an unbalanced panel. Only those states for which data are available for all years have been included. The two exceptions are Maharashtra and Tamilnadu. Data for Bombay have been treated as earlier observations of the time series for Maharashtra, even though the two states are not coterminous. Additionally, time was included as an explanatory variable in the intercept equation to capture the technical change. The model with a Cobb-Douglas production function is

$$\begin{aligned} \ln(Y_{it}) &= \alpha_{it} + \beta_L \ln(L_{it}) + \beta_K \ln(K_{it}) + v_{it} \\ \alpha_{it} &= \alpha_i + \phi\alpha_{i,t-1} + \gamma \text{Time}_t + u_{it}; \quad u_{it} \leq 0. \end{aligned} \tag{2.35}$$

If we define

$$\begin{aligned} y_{it} &= \ln(Y_{it}), \\ l_{it} &= \ln(L_{it}), \text{ and} \\ k_{it} &= \ln(K_{it}) \end{aligned} \tag{2.36}$$

then the estimated model becomes

$$\begin{aligned} y_{it} &= \tilde{\alpha}_i + \phi y_{i,t-1} + \beta_L l_{it} + \beta_K k_{it} - \phi\beta_L l_{i,t-1} - \phi\beta_K k_{i,t-1}\beta' + \gamma \text{ Time}_t + \tilde{\varepsilon}_{it} \\ &\text{where } \tilde{\varepsilon}_{it} = (v_{it} - \phi v_{it-1} + \tilde{u}_{it}), \\ &\qquad u_{it} = \tilde{u}_{it} - \max_t\{\tilde{u}_{it}\} \leq 0, \\ &\qquad \alpha_i = \tilde{\alpha}_i + \max_t\{\tilde{u}_{it}\}. \end{aligned} \tag{2.37}$$

First we estimate a Cobb-Douglas production function using Ordinary least squares, and the results are presented in Table 2.2. The elasticity of output with respect to Labor is 74.45% and the elasticity with respect of Capital is 24.39%. We used Ordinary Least Squares estimates of parameters as the initial values for the estimation of the model in (2.32). Use of the Newton-Raphson optimization method led to the estimates reported in Table 2.3. We find that the elasticity of Labor is 0.81 and the elasticity of Capital is 0.29. The coefficient of time that captures the technological change is 0.001. We apply the Greene correction by adding the maximum of the estimated error for each firm to the estimated firm specific intercept. We appropriately adjust the error terms and we use (2.16) to estimate the technical efficiency across firms and time. As we can see from the results the state of MP has the higher intercept and thus its frontier lies higher than the other states. The estimates of the technical efficiency are reported in Table 2.4. The overall technical efficiency for the 12 states of our sample during the period 1954-1983 is 86.23%. When a state has efficiency equal to 1, this implies that this state was 100% efficient at that point of time. Technical Efficiency values less than 1 imply that the state was inefficient. From the state average Technical Efficiency we can see that the state of Bihar has 93.15% efficiency over the period that we study and it is the state in our sample that is closest to its frontier. Notice that the frontier of the state of Bihar is ranked fifth relative to the other states. The state of MP that has the highest located frontier intercept is ranked last in terms of Technical Efficiency with only 65.91% efficiency.

2.5 Conclusion

Frontier production function models with time varying technical efficiency typically model a firm's efficiency level as a systematic function of time. As a result, one can not distinguish between technical change and changes in the level of efficiency. In this paper, we model efficiency change through unit-specific intercepts that evolve over time like a first order auto regressive (AR(1)) process. This builds on the Cooley-Prescott adaptive regression model within the class of regression models with time-varying parameters. The current approach has two major advantages. The first is that it allows past levels of inefficiency to influence the current position of the firm's frontier though the intercept. The second advantage is that it is possible to estimate both the technical change and efficiency change.

Table 2.1: List of Indian States	
State	**Number**
Assam	1
Bihar	2
Delhi	3
Himachal Pradesh	4
Madhya Pradesh	5
Orissa	6
Punjab	7
Rajasthan	8
Uttar Pradesh	9
West Bengal	10
Andhra Pradesh	11
Maharashtra	12

Table 2.2: Regression results

	Variable	Parameter Estimate	Standard Error	T statistic for $\vartheta=0$
Assam	α_1	-2.58885	0.41176	-6.287
Bihar	α_2	-2.29305	0.45486	-5.041
Delhi	α_3	-2.40706	0.42257	-5.696
Himachal Pradesh	α_4	-2.38347	0.33000	-7.223
Madhya Pradesh	α_5	-2.62344	0.44192	-5.936
Orissa	α_6	-2.52943	0.40089	-6.31
Punjab	α_7	-2.52923	0.42629	-5.933
Rajasthan	α_8	-2.63343	0.41627	-6.326
Uttar Pradesh	α_9	-2.63640	0.47900	-5.504
West Bengal	α_{10}	-2.42462	0.51483	-4.71
Andhra Pradesh	α_{11}	-2.81939	0.46600	-6.05
Maharashtra	α_{12}	-2.09154	0.51431	-4.067
ln(Labor)	β_L	0.74451	0.06485	11.48
ln(Capital)	β_K	0.24387	0.04683	5.208
time	γ	0.00978	0.00349	2.803
adjusted $R^2 = 0.9991$				

Table 2.3: Maximum Likelihood Estimates

Firm Specific Intercepts	$\tilde{\hat{\alpha}}_i$	Parameter Estimate	$\max_t\{\hat{u}_{it}\}$	$\hat{\alpha}_i = \tilde{\hat{\alpha}}_i + \max_t\{\hat{u}_{it}\}$	Standard Error	t-statistic	Relative Rank of the intercepts
Assam	α_1	-2.29607	0.24074	-2.05533	0.33067	-6.21560	3
Bihar	α_2	-2.14388	0.07818	-2.06570	0.41974	-4.92140	5
Delhi	α_3	-2.17649	0.11753	-2.05896	0.36746	-5.60329	4
Himachal Pradesh	α_4	-2.14404	0.13284	-2.01121	0.33468	-6.00934	2
Madhya Pradesh	α_5	-1.95888	0.44879	-1.51010	0.25991	-5.81003	1
Orissa	α_6	-2.33023	0.17227	-2.15796	0.35251	-6.12170	7
Punjab	α_7	-2.23765	0.16137	-2.07628	0.32143	-6.45944	6
Rajasthan	α_8	-2.25180	0.07957	-2.17223	0.34096	-6.37101	8
Uttar Pradesh	α_9	-2.28736	0.09276	-2.19460	0.33215	-6.60731	9
West Bengal	α_{10}	-2.38909	0.13021	-2.25888	0.38235	-5.90786	11
Andhra Pradesh	α_{11}	-2.31304	0.07439	-2.23865	0.41399	-5.40743	10
Maharashtra	α_{12}	-2.44249	0.12783	-2.31466	0.36920	-6.26947	12
	Varia-ble	**Parameter Estimate**	**Standard Error**	**t-statistic**			
ln(Labor)	β_L	0.81026	0.07869	10.29692			
ln(Capital)	β_K	0.29442	0.05176	5.68866			

time	γ	0.00133	0.00027	4.76320
	ϕ	0.38465	0.02611	14.73107
	θ	0.26655	0.05444	4.89651
	σ_w^2	0.04774	0.00341	14.01577
	σ_v^2	0.03308		
	σ_u^2	0.01315		

Table 2.4: Technical Efficiency

State \ Year	1	2	3	4	5	6	7	8	9	10	11	12	Average TE across States
1955	0.801	0.904	0.971	0.900	0.669	0.812	0.903	0.917	0.953	0.884	0.949	0.882	0.879
1956	0.874	0.962	0.995	1.000	0.655	0.887	0.951	0.990	0.951	0.929	0.969	1.000	0.930
1957	0.821	0.888	0.977	0.943	0.723	0.901	0.857	0.962	0.761	0.867	0.934	0.846	0.873
1958	0.801	0.875	0.928	0.931	0.843	0.814	0.841	0.950	0.863	0.885	0.932	0.858	0.877
1959	0.808	0.919	0.879	0.928	0.793	0.801	0.856	1.000	0.909	0.879	0.971	0.905	0.887
1960	1.000	0.937	0.939	0.970	1.000	0.807	1.000	0.913	0.850	0.891	0.948	0.826	0.924
1961	0.821	0.912	0.835	0.931	0.634	0.884	0.824	0.917	0.910	0.904	0.917	0.846	0.861
1962	0.903	0.923	0.873	0.941	0.825	0.884	0.962	0.953	0.948	0.881	0.914	0.891	0.908
1963	0.801	0.937	0.923	0.894	0.303	0.574	0.705	0.917	0.917	0.873	0.929	0.835	0.801
1964	0.774	0.926	0.880	0.848	0.616	0.840	0.888	0.984	0.908	0.849	0.930	0.832	0.856
1965	0.757	0.911	0.911	0.879	0.579	0.794	0.841	0.863	0.834	0.883	0.924	0.835	0.834
1966	0.712	0.921	0.876	0.872	0.487	0.774	0.888	0.908	0.882	0.901	0.897	0.844	0.830
1967	0.819	0.908	0.889	0.827	0.546	0.788	0.733	0.903	0.854	0.832	0.864	0.822	0.815
1968	0.702	0.894	0.804	0.809	0.496	0.765	0.652	0.883	0.908	0.764	0.856	0.776	0.776
1969	0.645	0.899	0.856	0.750	0.789	0.859	0.864	0.930	0.895	0.952	0.886	0.773	0.841
1970	0.776	0.943	0.851	0.922	0.579	0.779	0.788	0.917	0.893	0.918	0.885	0.878	0.844
1971	0.792	0.950	0.869	0.845	0.718	0.848	0.872	0.906	0.927	0.862	0.896	0.866	0.863
1972	0.732	0.955	0.863	0.867	0.649	0.828	0.779	0.947	0.933	0.824	0.921	0.876	0.848

1974	0.867	1.000	0.891	0.834	0.486	1.000	0.963	0.896	0.951	0.939	0.989	0.969	0.899
1975	0.899	0.998	1.000	0.893	0.526	0.949	0.883	0.966	1.000	0.902	1.000	0.913	0.911
1976	0.830	0.942	0.893	0.782	0.697	0.834	0.824	0.947	0.941	0.850	0.936	0.930	0.867
1977	0.866	0.937	0.821	0.871	0.684	0.908	0.953	0.855	0.935	0.872	0.934	0.890	0.877
1978	0.860	0.950	0.774	0.856	0.700	0.855	0.840	0.936	0.945	0.793	0.919	0.872	0.858
1979	0.775	0.949	0.838	0.846	0.657	0.819	0.900	0.887	0.966	0.824	0.927	0.869	0.855
1980	0.782	0.936	0.804	0.832	0.701	0.863	0.915	0.907	0.973	0.814	0.920	0.861	0.859
1981	0.694	0.954	0.767	0.831	0.682	0.919	0.791	0.936	0.941	0.803	0.945	0.875	0.845
1982	0.764	0.938	0.961	0.842	0.706	0.921	0.821	0.874	0.913	1.000	0.903	0.873	0.876
1983	0.753	0.912	0.951	0.831	0.714	0.889	0.764	0.851	0.862	0.843	0.902	0.942	0.851
State Average TE													
State #	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**Overall Average**
	0.801	0.931	0.886	0.874	0.659	0.843	0.852	0.922	0.912	0.872	0.925	0.871	0.862
rank	11	1	5	6	12	10	9	3	4	7	2	8	

Table 2.5: Data

State	Year	Labor	Capital	Value Added
Andhra Pradesh	1954	39430	2890.31	930.75
Assam	1954	6687	694.7	262.73
Bihar	1954	111021	17575.87	7841.14
Delhi	1954	24393	1725.78	1354.38
Himachal Pradesh	1954	765	135.28	20.37
Madhya Pradesh	1954	52234	3979.89	1458.25
Maharashtra	1954	540880	42685.56	25608.96
Orissa	1954	16152	1919.56	706.72
Punjab	1954	27183	2559.41	977.6
Rajasthan	1954	13751	2129.8	409.37
Uttar Pradesh	1954	170560	12228.52	5753.56
West Bengal	1954	452353	30127.97	16468.43
Andhra Pradesh	1955	41365	3394.11	1095.83
Assam	1955	6877	740.33	247.92
Bihar	1955	111207	17688.77	8618.75
Delhi	1955	26286	1941.07	1314.58
Himachal Pradesh	1955	710	127.07	31.25
Madhya Pradesh	1955	51150	4186	1477.08
Maharashtra	1955	538223	43438.31	27414.58

Orissa	1955	16443	1797.42	758.33
Punjab	1955	27905	2937.38	1047.92
Rajasthan	1955	15459	2174.95	589.58
Uttar Pradesh	1955	172859	13462.25	6068.75
West Bengal	1955	465357	31473.3	19441.67
Andhra Pradesh	1956	42373	3401.08	1582.96
Assam	1956	8270	751.8	347.53
Bihar	1956	116455	20266.19	10280.27
Delhi	1956	28019	2109.71	1766.82
Himachal Pradesh	1956	727	116.91	35.87
Madhya Pradesh	1956	52306	4007.19	1867.71
Maharashtra	1956	551236	42197.84	33318.39
Orissa	1956	16266	1991.01	910.31
Punjab	1956	28888	3253.6	1352.02
Rajasthan	1956	16530	2258.99	695.07
Uttar Pradesh	1956	170057	16640.29	7260.09
West Bengal	1956	487753	33868.71	22903.59
Andhra Pradesh	1957	68910	6622.57	2148.68
Assam	1957	8632	832.45	340.12
Bihar	1957	122951	22467.37	10981.67
Delhi	1957	30621	2313.93	1832.99
Himachal Pradesh	1957	807	118.17	52.95

Madhya Pradesh	1957	20239	1516.75	739.31
Maharashtra	1957	626822	53601.41	34883.91
Orissa	1957	17493	2153.44	810.59
Punjab	1957	39394	4243.39	1890.02
Rajasthan	1957	25653	2862.43	631.36
Uttar Pradesh	1957	181980	17772.49	6930.75
West Bengal	1957	494600	38481.48	22855.4
Andhra Pradesh	1958	70300	7032.48	2066.28
Assam	1958	7895	837.61	288.5
Bihar	1958	125587	25897.44	10214.42
Delhi	1958	31593	2567.52	1836.26
Himachal Pradesh	1958	898	135.04	99.42
Madhya Pradesh	1958	58605	5885.47	2128.65
Maharashtra	1958	591319	56162.39	30931.77
Orissa	1958	19641	2552.14	828.46
Punjab	1958	40579	4418.8	1902.53
Rajasthan	1958	24798	2796.58	686.16
Uttar Pradesh	1958	186525	18164.1	7161.79
West Bengal	1958	476769	38452.99	21643.27
Andhra Pradesh	1959	65348	7284.99	2192.31
Assam	1959	6658	713.32	236.54
Bihar	1959	125582	35263.07	9300

Delhi	1959	32462	2762.23	1867.31
Himachal Pradesh	1959	995	138.28	113.46
Madhya Pradesh	1959	37360	4406.41	1207.69
Maharashtra	1959	559589	56765.6	32496.15
Orissa	1959	19366	3055.65	861.54
Punjab	1959	41974	5268.13	2321.15
Rajasthan	1959	21022	2738.62	711.54
Uttar Pradesh	1959	172448	16457	6430.77
West Bengal	1959	478630	40994.94	24317.31
Andhra Pradesh	1960	137000	9475.41	3580.71
Assam	1960	69000	11065.57	6011.13
Bihar	1960	174000	40360.66	13896.1
Delhi	1960	40000	3147.54	2541.74
Himachal Pradesh	1960	1000	229.51	241.19
Madhya Pradesh	1960	87000	5196.72	2448.98
Maharashtra	1960	607000	67393.44	39814.47
Orissa	1960	20000	6213.11	1614.1
Punjab	1960	69000	6754.1	3116.88
Rajasthan	1960	43000	3377.05	1224.49
Uttar Pradesh	1960	231000	19016.39	8775.51
West Bengal	1960	662000	60836.07	34805.19
Andhra Pradesh	1961	143000	12539.43	3982.91

Assam	1961	68000	10488.96	4376.07
Bihar	1961	166000	40283.91	10341.88
Delhi	1961	40000	3028.39	2358.97
Himachal Pradesh	1961	2000	236.59	170.94
Madhya Pradesh	1961	86000	9100.95	3572.65
Maharashtra	1961	626000	74763.41	40735.04
Orissa	1961	26000	6577.29	1452.99
Punjab	1961	68000	6324.92	2854.7
Rajasthan	1961	44000	3895.9	1487.18
Uttar Pradesh	1961	231000	21167.19	9521.37
West Bengal	1961	665000	68643.53	33008.55
Andhra Pradesh	1962	142000	13139.36	4561.98
Assam	1962	61000	10076.57	4512.4
Bihar	1962	176000	41638.59	11107.44
Delhi	1962	42000	3935.68	2661.16
Himachal Pradesh	1962	2000	229.71	231.4
Madhya Pradesh	1962	102000	10903.52	4710.74
Maharashtra	1962	650000	82266.46	44033.06
Orissa	1962	25000	8897.4	2000
Punjab	1962	76000	11332.31	4000
Rajasthan	1962	44000	3889.74	1735.54
Uttar Pradesh	1962	252000	25666.16	10347.11

West Bengal	1962	679000	79540.58	33619.83
Andhra Pradesh	1963	156000	14375	4513.56
Assam	1963	67000	13997.09	4178.63
Bihar	1963	186000	42136.63	13492.82
Delhi	1963	45000	4505.81	2583.73
Himachal Pradesh	1963	3000	305.23	31.9
Madhya Pradesh	1963	137000	45566.86	3189.79
Maharashtra	1963	674000	94375	49377.99
Orissa	1963	53000	40247.09	2902.71
Punjab	1963	86000	13197.67	4322.17
Rajasthan	1963	49000	5101.74	2009.57
Uttar Pradesh	1963	255000	25261.63	9984.05
West Bengal	1963	729000	109941.9	40191.39
Andhra Pradesh	1964	172000	27175.14	5587.79
Assam	1964	69000	13474.58	3526.72
Bihar	1964	196000	47175.14	13679.39
Delhi	1964	58000	9180.79	3267.18
Himachal Pradesh	1964	2000	169.49	45.8
Madhya Pradesh	1964	133000	51341.81	5618.32
Maharashtra	1964	691000	110423.7	52152.67
Orissa	1964	51000	42401.13	3938.93
Punjab	1964	95000	15466.1	5725.19

Rajasthan	1964	54000	6920.9	2305.34
Uttar Pradesh	1964	270000	44872.88	11389.31
West Bengal	1964	794000	123107.3	45251.91
Andhra Pradesh	1965	209000	27844.47	6541.24
Assam	1965	76000	16221.01	3560.06
Bihar	1965	202000	50409.28	15238.78
Delhi	1965	64000	9290.59	3632.42
Himachal Pradesh	1965	3000	450.2	101.3
Madhya Pradesh	1965	155000	59918.14	7163.53
Maharashtra	1965	755000	124665.8	55615.05
Orissa	1965	62000	43956.34	4602.03
Punjab	1965	134000	35852.66	7395.08
Rajasthan	1965	65000	14065.48	2633.86
Uttar Pradesh	1965	276000	51991.81	12952.24
West Bengal	1965	825000	138704	47814.76
Andhra Pradesh	1966	218000	29987.08	7083.91
Assam	1966	77000	18372.09	3053.65
Bihar	1966	212000	71085.27	16327.37
Delhi	1966	65000	10206.72	3741.4
Himachal Pradesh	1966	5000	762.27	123.8
Madhya Pradesh	1966	146000	68397.93	6850.07
Maharashtra	1966	767000	135762.3	58500.69

Orissa	1966	71000	47441.86	6011
Punjab	1966	139000	45219.64	8363.14
Rajasthan	1966	70000	15710.59	3053.65
Uttar Pradesh	1966	296000	63617.57	15749.66
West Bengal	1966	880000	161240.3	48762.04
Andhra Pradesh	1967	223000	41093.56	7487.62
Assam	1967	77000	19696.23	4108.91
Bihar	1967	211000	74835.97	16670.79
Delhi	1967	67000	10789.79	3403.47
Himachal Pradesh	1967	8000	1324.42	247.52
Madhya Pradesh	1967	156000	76852.98	7722.77
Maharashtra	1967	751000	151701.1	57710.4
Orissa	1967	72000	49951.4	4009.9
Punjab	1967	83000	32029.16	4962.87
Rajasthan	1967	72000	18201.7	3081.68
Uttar Pradesh	1967	284000	74009.72	13452.97
West Bengal	1967	864000	171518.8	43254.95
Andhra Pradesh	1968	225000	48605.99	6721.13
Assam	1968	74000	17453.92	2897.6
Bihar	1968	214000	84251.15	13616.56
Delhi	1968	69000	10426.27	3104.58
Himachal Pradesh	1968	9000	1152.07	217.86

Madhya Pradesh	1968	163000	77004.61	7527.23
Maharashtra	1968	746000	154804.2	54880.17
Orissa	1968	67000	54354.84	2450.98
Punjab	1968	72000	21866.36	3736.38
Rajasthan	1968	69000	20841.01	3540.31
Uttar Pradesh	1968	267000	77292.63	9814.81
West Bengal	1968	822000	167477	38529.41
Andhra Pradesh	1969	232000	52923.78	6584.05
Assam	1969	71000	18452.79	2068.97
Bihar	1969	216000	74254.84	14084.05
Delhi	1969	83000	19124	3426.72
Himachal Pradesh	1969	9000	1547.21	711.21
Madhya Pradesh	1969	167000	75460.75	10010.78
Maharashtra	1969	787000	172707.6	59008.62
Orissa	1969	74000	55233.22	4547.41
Punjab	1969	78000	25130.83	4644.4
Rajasthan	1969	70000	27303.75	3933.19
Uttar Pradesh	1969	272000	86075.09	16077.59
West Bengal	1969	754000	148111.5	36907.33
Andhra Pradesh	1970	265000	58193.04	10115.18
Assam	1970	75000	23793.49	3287.96
Bihar	1970	226000	85914.7	15445.03

Delhi	1970	74000	13187.43	4314.14
Himachal Pradesh	1970	11000	2300.79	659.69
Madhya Pradesh	1970	169000	72727.27	8680.63
Maharashtra	1970	811000	192873.2	70471.2
Orissa	1970	85000	52783.39	4900.52
Punjab	1970	78000	28204.26	4837.7
Rajasthan	1970	77000	29842.87	4376.96
Uttar Pradesh	1970	288000	101358	19193.72
West Bengal	1970	770000	166071.8	39623.04
Andhra Pradesh	1971	265000	57407.79	10901.72
Assam	1971	74000	24928.28	3961.5
Bihar	1971	235000	85788.93	16899.7
Delhi	1971	78000	13084.02	4265.45
Himachal Pradesh	1971	14000	2592.21	1246.2
Madhya Pradesh	1971	173000	67418.03	10151.98
Maharashtra	1971	836000	201280.7	78176.29
Orissa	1971	78000	49651.64	5744.68
Punjab	1971	84000	29774.59	5096.25
Rajasthan	1971	87000	32827.87	5460.99
Uttar Pradesh	1971	317000	115594.3	19037.49
West Bengal	1971	770000	150215.2	40091.19
Andhra Pradesh	1972	283000	58854.67	12296.37

Assam	1972	79000	24427.33	3660.45
Bihar	1972	246000	89499.52	17860.65
Delhi	1972	85000	9181.91	4308.15
Himachal Pradesh	1972	16000	4302.21	1462.22
Madhya Pradesh	1972	182000	64773.82	10500.49
Maharashtra	1972	899000	204533.2	86722.28
Orissa	1972	82000	46997.11	4877.33
Punjab	1972	92000	32406.16	6221.79
Rajasthan	1972	93000	35640.04	6241.41
Uttar Pradesh	1972	340000	121838.3	17419.04
West Bengal	1972	801000	152945.1	45338.57
Andhra Pradesh	1974	291000	58657.6	16632.56
Assam	1974	78000	23305.01	5237.21
Bihar	1974	236000	88105.35	18725.58
Delhi	1974	83000	17621.07	4734.88
Himachal Pradesh	1974	21000	4876.81	883.72
Madhya Pradesh	1974	203000	79056.92	19618.6
Maharashtra	1974	929000	227102.8	104930.2
Orissa	1974	86000	44052.68	7906.98
Punjab	1974	97000	39328.8	6241.86
Rajasthan	1974	99000	35063.72	7013.95
Uttar Pradesh	1974	356000	135794.4	24446.51

West Bengal	1974	797000	125616	52911.63
Andhra Pradesh	1975	327000	57883.31	17463.71
Assam	1975	98000	22042.06	7876.24
Bihar	1975	267000	82876.53	28120.7
Delhi	1975	75000	15976.93	4896.87
Himachal Pradesh	1975	20000	2455.9	634.07
Madhya Pradesh	1975	212000	73378.56	20954.93
Maharashtra	1975	969000	205909.1	110748.7
Orissa	1975	85000	47333.79	7708.17
Punjab	1975	102000	39857.53	7608.86
Rajasthan	1975	98000	35739.48	8174.18
Uttar Pradesh	1975	414000	127401.6	27249.81
West Bengal	1975	771000	110685.2	55477.46
Andhra Pradesh	1976	349000	56441.07	18873.7
Assam	1976	98000	21015.17	6959.64
Bihar	1976	307000	137666.3	31230.47
Delhi	1976	83000	19002.33	4244.79
Himachal Pradesh	1976	16000	3395.57	1139.32
Madhya Pradesh	1976	236000	78868.14	17272.14
Maharashtra	1976	947000	179620.8	92057.29
Orissa	1976	92000	34492.42	6217.45
Punjab	1976	106000	41901.98	8053.39

Rajasthan	1976	114000	46691.95	9127.6
Uttar Pradesh	1976	442000	120408.4	24375
West Bengal	1976	813000	110525.1	51276.04
Andhra Pradesh	1977	377000	63341.11	18876.4
Assam	1977	101000	23300.29	7865.17
Bihar	1977	309000	153586	24701.36
Delhi	1977	89000	22320.7	5552.93
Himachal Pradesh	1977	18000	5673.47	1803.67
Madhya Pradesh	1977	233000	88413.99	19520.99
Maharashtra	1977	993000	214658.9	94991.13
Orissa	1977	95000	42769.68	9130.69
Punjab	1977	127000	50979.59	7687.76
Rajasthan	1977	124000	45871.72	9243.05
Uttar Pradesh	1977	508000	140793	28799.53
West Bengal	1977	789000	119562.7	48710.82
Andhra Pradesh	1978	399000	75967.37	19091.92
Assam	1978	112000	32162	9487.47
Bihar	1978	333000	161299.5	20729.81
Delhi	1978	86000	22954.55	5559.89
Himachal Pradesh	1978	17000	6008.16	2022.28
Madhya Pradesh	1978	236000	96381.12	18055.71
Maharashtra	1978	1001000	239155	99966.57

Orissa	1978	99000	46497.67	7894.15
Punjab	1978	134000	59213.29	9565.46
Rajasthan	1978	123000	51800.7	9587.74
Uttar Pradesh	1978	561000	157237.8	25576.6
West Bengal	1978	845000	122039.6	48740.95
Andhra Pradesh	1979	431000	86266.74	20241.74
Assam	1979	111000	31679.69	7252.1
Bihar	1979	326000	207288	24341.39
Delhi	1979	91000	22633.93	5688.21
Himachal Pradesh	1979	18000	8989.96	2146.03
Madhya Pradesh	1979	245000	100128.4	16196.35
Maharashtra	1979	1044000	258616.1	106013.8
Orissa	1979	97000	49023.44	8697.58
Punjab	1979	158000	64224.33	10315.74
Rajasthan	1979	125000	57963.17	10666.01
Uttar Pradesh	1979	554000	171104.9	26147.02
West Bengal	1979	834000	131668.5	49309.32
Andhra Pradesh	1980	467000	94203.82	21350.74
Assam	1980	109000	25654.09	6235.93
Bihar	1980	310000	218980.9	22336.79
Delhi	1980	95000	27138.66	5925.26
Himachal Pradesh	1980	20000	9632.53	2701.49

Madhya Pradesh	1980	255000	107809.9	18469.16
Maharashtra	1980	1076000	277672.7	107483.1
Orissa	1980	100000	50465.46	9725.35
Punjab	1980	181000	76834.88	12449.35
Rajasthan	1980	150000	60744.73	13075.19
Uttar Pradesh	1980	609000	167800.1	27294.01
West Bengal	1980	867000	130842.7	49991
Andhra Pradesh	1981	448000	96246.7	21376.66
Assam	1981	107000	23828.19	4335.19
Bihar	1981	347000	186590.3	20226.59
Delhi	1981	95000	26185.02	5746.05
Himachal Pradesh	1981	19000	9290.75	2513.89
Madhya Pradesh	1981	267000	125740.1	24181.27
Maharashtra	1981	1072000	289237.9	112646.4
Orissa	1981	117000	50189.43	7905.09
Punjab	1981	168000	76101.32	12796.07
Rajasthan	1981	150000	68334.8	12672.08
Uttar Pradesh	1981	678000	172903.1	28755.88
West Bengal	1981	854000	153405.3	55061.99
Andhra Pradesh	1982	466000	95555.11	22184.72
Assam	1982	102000	20671.76	4495.55
Bihar	1982	326000	197216.4	32685.46

Delhi	1982	82000	25414.32	5192.88
Himachal Pradesh	1982	22000	11074.01	3230.71
Madhya Pradesh	1982	274000	138644.4	27500
Maharashtra	1982	1080000	313024.9	112852.4
Orissa	1982	119000	55547.06	7971.07
Punjab	1982	186000	79975.86	12295.99
Rajasthan	1982	156000	70547.06	11947.33
Uttar Pradesh	1982	692000	200341.9	52266.32
West Bengal	1982	820000	164167.3	49721.81
Andhra Pradesh	1983	490000	106213.2	28979.39
Assam	1983	101000	19482.24	4417.26
Bihar	1983	327000	215309.9	40022.54
Delhi	1983	101000	32932.73	6493.88
Himachal Pradesh	1983	24000	18484.5	4278.82
Madhya Pradesh	1983	294000	164130.8	28937.54
Maharashtra	1983	1053000	320589.6	102714.1
Orissa	1983	122000	61591.08	6918.87
Punjab	1983	194000	92588.81	11793.3
Rajasthan	1983	180000	77611.49	11506.76
Uttar Pradesh	1983	718000	232180.7	45035.42
West Bengal	1983	853000	162664.4	49439.79

APPENDIX 2.1 : The Variance-Covariance Matrix.

- The variance-covariance matrix Σ can be written as

$$\Sigma = E(\varepsilon_i \varepsilon_i^{'}) = \begin{bmatrix} (1+\theta^2)\sigma_w^2 & -\theta\sigma_w^2 & \cdots & 0 \\ -\theta\sigma_w^2 & (1+\theta^2)\sigma_w^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & -\theta\sigma_w^2 \\ 0 & 0 & -\theta\sigma_w^2 & (1+\theta^2)\sigma_w^2 \end{bmatrix} = \sigma_w^2\Omega, \tag{2.*i4}$$

where $\Omega = \begin{bmatrix} 1+\theta^2 & -\theta & \cdots & 0 \\ -\theta & 1+\theta^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & -\theta \\ 0 & 0 & -\theta & 1+\theta^2 \end{bmatrix}$.

From Hamilton (1994) we get the determinant of Ω:

$$|\Omega| = 1+\theta^2+\ldots+\theta^{2T} = \frac{1-\theta^{2(T+1)}}{1-\theta^2}. \tag{2.*i5}$$

- Inverse and determinant of Σ

$$\Sigma^{-1} = \frac{1}{\sigma_w^2}\Omega^{-1} \tag{2.*i5}$$

$$|\Sigma| = (\sigma_w^2)^T\,|\Omega| \tag{2.*i5}$$

- Derivatives of Ω

$$\frac{\partial\Omega}{\partial\theta} = \begin{bmatrix} 2\theta & -1 & \cdots & 0 \\ -1 & 2\theta & \cdots & \vdots \\ \vdots & \vdots & \ddots & -1 \\ 0 & 0 & \cdots & 2\theta \end{bmatrix}$$

$$\frac{\partial^2 \Omega}{\partial \theta^2} = 2\,\mathrm{I}\text{, where I is the identity TxT matrix}$$

- Derivatives of ln|Ω|

$$\frac{\partial \ln|\Omega|}{\partial \theta} = \mathrm{trace}(\Omega^{-1}\frac{\partial \Omega}{\partial \theta})$$

$$\frac{\partial^2 \ln|\Omega|}{\partial \theta^2} = \mathrm{trace}(-\Omega^{-1}\frac{\partial \Omega}{\partial \theta}\Omega^{-1}\frac{\partial \Omega}{\partial \theta} + 2\Omega^{-1}\mathrm{I}) = \mathrm{trace}(-\Omega^{-1}\frac{\partial \Omega}{\partial \theta}\Omega^{-1}\frac{\partial \Omega}{\partial \theta} + 2\Omega^{-1})$$

- Derivatives of Ω^{-1}

$$\frac{\partial \Omega^{-1}}{\partial \theta} = -\Omega^{-1}\frac{\partial \Omega}{\partial \phi}\Omega^{-1}$$

$$\frac{\partial^2 \Omega^{-1}}{\partial \theta^2} = +2\Omega^{-1}\frac{\partial \Omega}{\partial \phi}\Omega^{-1}\frac{\partial \Omega}{\partial \phi}\Omega^{-1} - 2\Omega^{-1}$$

- Useful Formulas:

$$\frac{\partial \ln|A|}{\partial \alpha} = \mathrm{tr}(A^{-1}\frac{\partial A}{\partial \alpha}) = \mathrm{tr}(A\frac{\partial A^{-1}}{\partial \alpha})$$ [Dhrymes (1978), p. 534] (2.*i1)

$$\frac{\partial A^{-1}}{\partial \alpha} = -A^{-1}\frac{\partial A}{\partial \alpha}A^{-1}$$ [Dhrymes (1978), p. 540] (2.*i2)

APPENDIX 2.2 : $\tilde{u}_{it} \mid \tilde{\varepsilon}_{it}$.

- $\tilde{u}_{it} \mid \tilde{\varepsilon}_{it} \sim N(\mu^*_{it}, \sigma^2_*)$

Recall that

$$v_{it} \sim \text{iid } N(0,\sigma_v^2) \Rightarrow v_{it} - \phi v_{it-1} \sim N(0,(1+\phi^2)\sigma_v^2),$$
$$\tilde{u}_{it} \sim \text{iid } N(0,\sigma_u^2), \text{ and}$$
$$\tilde{\varepsilon}_{it} = (v_{it} - \phi v_{it-1}) + \tilde{u}_{it} \sim N(0,(1+\phi^2)\sigma_v^2 + \sigma_u^2)$$

Now the probability density function (pdf) of $(v_{it} - \phi v_{it-1})$, $\tilde{u}_{it}$, and $\tilde{\varepsilon}_{it}$ are respectively

$$f(v_{it} - \phi v_{it-1}) = \frac{1}{\sqrt{2\pi}\sqrt{(1+\phi^2)\sigma_v^2}} \exp\{-\frac{(v_{it} - \phi v_{it-1})^2}{2(1+\phi^2)\sigma_v^2}\},$$

$$f(\tilde{u}_{it}) = \frac{1}{\sqrt{2\pi}\sqrt{\sigma_u^2}} \exp\{-\frac{\tilde{u}_{it}^2}{2\sigma_u^2}\}, \text{ and}$$

$$f(\tilde{\varepsilon}_{it}) = \frac{1}{\sqrt{2\pi}\sqrt{(1+\phi^2)\sigma_v^2 + \sigma_u^2}} \exp\{-\frac{\tilde{\varepsilon}_{it}^2}{2((1+\phi^2)\sigma_v^2 + \sigma_u^2)}\}.$$

The joint pdf of $(v_{it} - \phi v_{it-1})$ and $\tilde{u}_{it}$ will be

$$f(v_{it} - \phi v_{it-1}, \tilde{u}_{it}) = \frac{1}{(\sqrt{2\pi})^2\sqrt{(1+\phi^2)\sigma_v^2\sigma_u^2}} \exp\{-\frac{(v_{it} - \phi v_{it-1})^2}{2(1+\phi^2)\sigma_v^2} - \frac{\tilde{u}_{it}^2}{2\sigma_u^2}\}.$$

Now the joint pdf of $\tilde{\varepsilon}_{it}$ and $\tilde{u}_{it}$ can be written as

$$f(\tilde{\varepsilon}_{it},\tilde{u}_{it})=\frac{1}{(\sqrt{2\pi})^2\sqrt{(1+\phi^2)\sigma_v^2\sigma_u^2}}\exp\{-\frac{(\tilde{\varepsilon}_{it}-\tilde{u}_{it})^2}{2(1+\phi^2)\sigma_v^2}-\frac{\tilde{u}_{it}{}^2}{2\sigma_u^2}\}$$

We can find that the exponential term is equal to

$$\frac{(\tilde{\varepsilon}_{it}-\tilde{u}_{it})^2}{(1+\phi^2)\sigma_v^2}+\frac{\tilde{u}_{it}{}^2}{\sigma_u^2}=\frac{\tilde{u}_{it}{}^2}{\sigma_u^2}+\frac{\tilde{u}_{it}{}^2}{(1+\phi^2)\sigma_v^2}-2\frac{\tilde{u}_{it}\tilde{\varepsilon}_{it}}{(1+\phi^2)\sigma_v^2}+\frac{\tilde{\varepsilon}_{it}{}^2}{(1+\phi^2)\sigma_v^2}=$$

$$=\frac{(1+\phi^2)\sigma_v^2+\sigma_u^2}{(1+\phi^2)\sigma_v^2\sigma_u^2}\tilde{u}_{it}{}^2-2\frac{\tilde{u}_{it}\tilde{\varepsilon}_{it}}{(1+\phi^2)\sigma_v^2}+\frac{\tilde{\varepsilon}_{it}{}^2}{(1+\phi^2)\sigma_v^2}=$$

$$=\frac{(1+\phi^2)\sigma_v^2+\sigma_u^2}{(1+\phi^2)\sigma_v^2\sigma_u^2}(\tilde{u}_{it}{}^2-2\tilde{u}_{it}\frac{\tilde{\varepsilon}_{it}\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2})+\frac{\tilde{\varepsilon}_{it}{}^2}{(1+\phi^2)\sigma_v^2}=$$

$$=\frac{(1+\phi^2)\sigma_v^2+\sigma_u^2}{(1+\phi^2)\sigma_v^2\sigma_u^2}(\tilde{u}_{it}-\frac{\tilde{\varepsilon}_{it}\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2})^2-$$

$$-\frac{\tilde{\varepsilon}_{it}{}^2\sigma_u^2}{[(1+\phi^2)\sigma_v^2+\sigma_u^2][(1+\phi^2)\sigma_v^2]}+\frac{\tilde{\varepsilon}_{it}{}^2}{(1+\phi^2)\sigma_v^2}=$$

$$=\frac{(1+\phi^2)\sigma_v^2+\sigma_u^2}{(1+\phi^2)\sigma_v^2\sigma_u^2}(\tilde{u}_{it}-\frac{\tilde{\varepsilon}_{it}\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2})^2-$$

$$-\frac{\tilde{\varepsilon}_{it}{}^2}{(1+\phi^2)\sigma_v^2}\left[\frac{\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}-1\right]=$$

$$=\frac{1}{\frac{(1+\phi^2)\sigma_v^2\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}}(\tilde{u}_{it}-\frac{\tilde{\varepsilon}_{it}\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2})^2-\frac{\tilde{\varepsilon}_{it}{}^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}$$

Now the joint pdf of $\tilde{\varepsilon}_{it}$ and $\tilde{u}_{it}$ can be further written as

$$f(\tilde{\varepsilon}_{it},\tilde{u}_{it})=\frac{1}{(\sqrt{2\pi})^2\sqrt{(1+\phi^2)\sigma_v^2\sigma_u^2}}\exp\{-\frac{1}{2\frac{(1+\phi^2)\sigma_v^2\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}}(\tilde{u}_{it}-\frac{\tilde{\varepsilon}_{it}\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2})^2\}$$

$$\exp\{\frac{\tilde{\varepsilon}_{it}{}^2}{2(1+\phi^2)\sigma_v^2+\sigma_u^2}\}$$

Using Bayes theorem, the conditional pdf of $\tilde{u}_{it}\mid\tilde{\varepsilon}_{it}$ is

$$f(\tilde{u}_{it}\mid\tilde{\varepsilon}_{it})=\frac{f(\tilde{\varepsilon}_{it},\tilde{u}_{it})}{f(\tilde{\varepsilon}_{it})}\Rightarrow$$

$$f(\tilde{u}_{it}\mid\tilde{\varepsilon}_{it})=\frac{1}{\sqrt{2\pi}\sqrt{\frac{(1+\phi^2)\sigma_v^2\sigma_u^2}{[(1+\phi^2)\sigma_v^2+\sigma_u^2]}}}\exp\{-\frac{1}{2\frac{(1+\phi^2)\sigma_v^2\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}}(\tilde{u}_{it}-\frac{\tilde{\varepsilon}_{it}\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2})^2\}\Rightarrow$$

$$f(\tilde{u}_{it}\mid\tilde{\varepsilon}_{it})=\frac{1}{\sqrt{2\pi}\sqrt{\sigma_*^2}}\exp\{-\frac{1}{2\sigma_*^2}(\tilde{u}_{it}-\mu_{it}^*)^2\},$$

where $\mu_{it}^*=\frac{\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}\tilde{\varepsilon}_{it}$ and $\sigma_*^2=\frac{(1+\phi^2)\sigma_v^2\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}$.

Thus the conditional distribution of $\tilde{u}_{it}\mid\tilde{\varepsilon}_{it}$ follows the normal distribution with mean μ_{it}^* and variance σ_*^2:

$$\tilde{u}_{it}\mid\tilde{\varepsilon}_{it}\sim N(\mu_{it}^*,\sigma_*^2),$$

where $\mu_{it}^*=\frac{\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}\tilde{\varepsilon}_{it}$ and $\sigma_*^2=\frac{(1+\phi^2)\sigma_v^2\sigma_u^2}{(1+\phi^2)\sigma_v^2+\sigma_u^2}$

- $E(e^{\tilde{u}_{it}} \mid \tilde{\varepsilon}_{it})$

$$E(e^{\tilde{u}_{it}} \mid \tilde{\varepsilon}_{it}) = \int_{-\infty}^{\infty} e^{\tilde{u}_{it}} f(\tilde{u}_{it} \mid \tilde{\varepsilon}_{it})\, d\tilde{u}_{it} =$$

$$= \frac{1}{\sqrt{2\pi\sigma_*^2}} \int_{-\infty}^{\infty} e^{\tilde{u}_{it}} e^{-(\tilde{u}_{it}-\mu_{it}^*)^2/(2\sigma_*^2)}\, d\tilde{u}_{it} =$$

$$= \frac{1}{\sqrt{2\pi\sigma_*^2}} \int_{-\infty}^{\infty} e^{-[(\tilde{u}_{it}-\mu_{it}^*)^2/(2\sigma_*^2)-\tilde{u}_{it}]}\, d\tilde{u}_{it}$$

$$\frac{(\tilde{u}_{it}-\mu_{it}^*)^2}{2\sigma_*^2} - \tilde{u}_{it} = \frac{\tilde{u}_{it}^2 - 2\tilde{u}_{it}\mu_{it}^* + \mu_{it}^{*2} - 2\tilde{u}_{it}\sigma_*^2}{2\sigma_*^2} = \frac{\tilde{u}_{it}^2 - 2\tilde{u}_{it}(\mu_{it}^* + \sigma_*^2) + \mu_{it}^{*2}}{2\sigma_*^2} =$$

$$= \frac{[\tilde{u}_{it}^2 - (\mu_{it}^* + \sigma_*^2)]^2 - (\mu_{it}^* + \sigma_*^2)^2 + \mu_{it}^{*2}}{2\sigma_*^2} =$$

$$= \frac{[\tilde{u}_{it}^2 - (\mu_{it}^* + \sigma_*^2)]^2}{2\sigma_*^2} - \frac{(2\mu_{it}^* + \sigma_*^2)\sigma_*^2}{2\sigma_*^2} =$$

$$= \frac{[\tilde{u}_{it}^2 - (\mu_{it}^* + \sigma_*^2)]^2}{2\sigma_*^2} - (\mu_{it}^* + \frac{\sigma_*^2}{2})$$

Finally,

$$E(e^{\tilde{u}_{it}} \mid \tilde{\varepsilon}_{it}) = \frac{1}{\sqrt{2\pi\sigma_*^2}} e^{(\mu_{it}^*+\sigma_*^2/2)} \int_{-\infty}^{\infty} e^{-(\tilde{u}_{it}-\mu_{it}^*-\sigma_*^2)^2/(2\sigma_*^2)}\, d\tilde{u}_{it} =$$

$$= e^{(\mu_{it}^*+\sigma_*^2/2)}$$

CHAPTER 3: MATHEMATICAL PROGRAMMING ESTIMATION OF A PARAMETRIC PRODUCTION FRONTIER

3.1 Introduction

Aigner and Chu (1968) developed a mathematical programming method for the estimation of a parametric production function with a one-sided error term to ensure that the estimated production function will exceed the output level actually produced from a given bundle of inputs. There are two problems with this approach. First, the estimated frontier is deterministic (non-stochastic) and does not allow any stochastic noise to influence the frontier. No account is taken of exogenous shocks over which the firm has no control and any deviation from the frontier is treated as technical inefficiency. The second limitation of this mathematical programming approach is that even though the input-output data set is only a sample from some underlying population, the sampling distribution of the estimated parameters cannot be derived analytically. As a result, we can not construct confidence intervals for the estimated parameters. As an ad hoc adjustment for the possibility of statistical noise, Timmer (1971) extended the Aigner and Chu approach by allowing an arbitrary percentage of observations to lie above the frontier. Richmond (1974) developed a method for the econometric estimation of a parametric deterministiv production frontier, by specifying a gamma distribution for the disturbance term. Greene (1990) proposed the corrected OLS (COLS) procedure, where the intercept is sufficiently adjusted to bring the observed data points below the frontier. But the resulting frontier also remains deterministic.

In a pair of seminal papers, Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) introduced the stochastic production frontier. The specified model incorporates a composed error term (ε_i), which is the sum of the exogenous shocks represented by a two-sided error term (v_i) and the technical efficiency that is represented by a one-sided error term (u_i):

$$\varepsilon_i = u_i - v_i. \tag{3.1}$$

The composed error term models are estimated by the maximum likelihood procedure, which requires that the statistical distributions of the components have to be explicitly specified. However, such specifications are generally arbitrary. Further, alternative distributional assumptions usually lead to different conclusions about the technical efficiency level of a firm (Greene, 1993). Another disadvantage of estimating a stochastic frontier using econometrics is that one cannot impose any inequality restrictions on the estimated coefficients based on the economic theory. For example, we cannot ensure that marginal productivities of inputs would be non-negative. Finally, one problem with any parametrically specified function is that the validity of any inference drawn from the fitted model is contingent on the validity of the specified form.

In the nonparametric approach known as Data Envelopment Analysis (DEA), introduced in the Operations Research literature by Charnes, Cooper, and Rhodes (1978, 1981) and further refined by Banker, Charnes, and Cooper (1984), one makes only a minimum number of regularity assumptions about the technology but leaves the exact form of a production, cost, or profit function unspecified. Because DEA also relies on mathematical programming, the resulting efficiency measures lack statistical properties. In a number of recent papers, Simar and Wilson (1992, 1995) have resorted to the bootstrap procedure in an effort to generate an empirical distribution function of the DEA efficiency measure. One can, therefore, construct confidence intervals from the empirical distribution.

A major drawback of the DEA procedure is that it cannot be used to predict the maximum output producible from any input bundle, which is not already observed in the sample. Thus, it is not useful for out-of-sample prediction. For this purpose, one must have a parametric function. Another disadvantage of the DEA methodology is that the estimated frontier is deterministic and it ignores any exogenous shocks that might influence the firm's behavior.

In this chapter, we revive the mathematical programming model by Aigner and Chu (1968), but we append a composed error term to a parametric frontier. As in the econometric models, the composed error is the sum of a two-sided random shock and a one-sided error representing inefficiency. Further, the mathematical programming model allows us to impose inequality restrictions on the coefficients.

The principal innovation in this paper lies in the fact that no assumptions are made about the distribution of the error terms except that they are independently distributed. Also, the proposed method allows us to impose inequality restrictions on the estimated parameters. It is well known that, when inequality restrictions are imposed, OLS or maximum likelihood estimation leads to not well defined statistical distributions of the estimated parameters. Hence, the resulting confidence intervals from restricted OLS estimation and the relevant test statistics might be invalid (Yancey 1981, and Judge et al, 1985). On the other hand, the estimated parameters obtained by any mathematical programming method are point estimates and have no statistical properties. We overcome this problem by applying bootstrap methods to obtain the statistical properties of the estimated stochastic frontier. This chapter is organized as following. In section 3.2, the quadratic programming model is laid out. Section 3.3 describes the smoothed bootstrap procedure applied. Section 3.4 includes an empirical application using state-level data on manufacturing output and inputs obtained from the 1992 Census of Manufactures.

3.2 The Quadratic Programming Model

Consider a data set (x^i, y_i) from n firms. Let y_i and x^i represent respectively the scalar output level and the input vector of k inputs for firm i in logarithmic terms. Assume that there is a monotonic frontier production function:

$$y_i^f = f(x^i;\beta) \quad i=1,2,\ldots,n, \tag{3.2}$$

where y_i^f is the maximum (or frontier) output obtainable from input bundle x^i and $\beta=\{\beta_1, \beta_2, \ldots, \beta_k\}$ is a vector of parameters to be estimated. In the stochastic specification the observed output is related to the unobservable frontier output as:

$$y_i = y_i^f + \varepsilon_i = f(x^i;\beta) + \varepsilon_i ; \quad \varepsilon_i = v_i - u_i ; \quad i=1,2,\ldots,n, \tag{3.3}$$

where v_i represents a two-sided error term reflecting random shifts in the frontier due to both favorable and unfavorable shocks and $u_i \geq 0$ is a one-sided error term that represents technical efficiency. We assume that both u_i and v_i are independently and identically distributed for any i=1,2,...,n. Further, v_i and u_j are statistically independent of each other i,j=1,2,...,n. Finally, we assume that the two-sided error terms v_i have zero mean. No other assumptions are made about the specific distribution of the error terms.

We can then estimate the above production function by solving the following quadratic programming (QP) problem:

$$\min \quad \frac{1}{2}c\sum_{i=1}^{n}v_i^2+\frac{1}{2}(1-c)\sum_{i=1}^{n}\left(u_i-\frac{1}{n}\sum_{j=1}^{n}u_j\right)^2$$

$$\text{st} \quad f(x^i;\beta)+v_i-u_i=y_i\,;\; i=1,2,\ldots,n$$

$$R(\beta)\le 0 \tag{3.4}$$

$$\sum_{i=1}^{n}v_i=0$$

$$u_i\ge 0;\;\; i=1,2,\ldots,n.$$

Here c is the weight $(0<c<1)$ assigned to the component due to exogenous shocks and $(1\text{-}c)$ is the weight assigned to the efficiency component. Further, $R(\beta)\le 0$ is a set of restrictions on the parameters. For the deterministic frontier model, similar to the Aigner and Chu (1968) frontier, we set v_i=0 for each i and the problem reduces to

$$\min \quad \frac{1}{2}\sum_{i=1}^{n}\left(u_i-\frac{1}{n}\sum_{j=1}^{n}u_j\right)^2$$

$$\text{st} \quad f(x^i;\beta)+v_i-u_i=y_i\,;\; i=1,2,\ldots,n$$

$$R(\beta)\le 0 \tag{3.5}$$

$$u_i\ge 0;\;\; i=1,2,\ldots,n.$$

Notice, that the first term of the objective function in (3.4) disappears when either c=0 or each v_i equals 0. However, when c=0, the v_i s feature in the restrictions and affect the optimal values of the u_i s.

On the other hand, when technical inefficiency is ignored, the problem becomes

$$\min \quad \frac{1}{2}\sum_{i=1}^{n}(v_i)^2$$

$$\text{st} \quad f(x^i;\beta)+v_i-u_i=y_i\,;\; i=1,2,\dots,n$$

$$R(\beta)\le 0 \tag{3.6}$$

$$\sum_{i=1}^{n}v_i=0.$$

As in the previous case, the second term of the objective function in (3.4) disappears when either c=1 or each u_i equals 0. However, when c=1, the u_i s feature in the restrictions and affect the optimal values of the v_i s.

In a stochastic frontier model, where both terms in the objective function are retained (0<c<1), when c< 0.5 then 1-c >0.5 and greater emphasis is given to the technical efficiency term, while when c>0.5 (1-c < 0.5), greater emphasis is given to the random shocks. We can vary these weights to examine the sensitivity in the distribution of the error terms. The term $R(\beta) \le 0$ refer to the set of restrictions that one might want to impose, such that the β_j s ($j=1,2,\dots k$) satisfy certain economic conditions.

If we assume that the production function is Cobb-Douglas and that inputs and outputs are expressed in natural log, then the previous quadratic programming problem takes the form:

$$\min \quad \frac{1}{2}c\sum_{i=1}^{n}v_i^2 + \frac{1}{2}(1-c)\sum_{i=1}^{n}\left(u_i - \frac{1}{n}\sum_{l=1}^{n}u_l\right)^2$$

$$\text{st} \quad \beta_0 + \sum_{j=1}^{K}\beta_j x_{ij} + v_i - u_i = y_i;\ i = 1,2,\ldots,n$$

$$R(\beta) \le 0 \qquad (3.7)$$

$$\sum_{i=1}^{n} v_i = 0$$

$$u_i \ge 0;\ \ i = 1,2,\ldots,n.$$

One problem in the econometric production frontier is often the marginal productivities (or equivalently, the partial elasticities) of inputs are negative. Even when we impose non-negativity restrictions in the QP-problem some of the β_j s ($j=1,2,\ldots,k$) may be zero at the optimal solution. To avoid this problem, we impose upper and lower bounds on them.

Note that under competitive conditions:

$$\beta_j = \frac{\partial \ln y}{\partial \ln x_j} = \frac{\partial y}{\partial x_j}\frac{x_j}{y} = \frac{w_j x_j}{py};\ j = 1,2,\ldots,k, \qquad (3.8)$$

where w_j is the price of j-th input and p is the price of the output. Thus, each β_j ($j=1,2,\ldots,k$) becomes the ratio of the expenditure on input j to the total output. When this information is available, as is the case in our empirical application, we can set the highest and the lowest observed values of this ratio as respectively, the upper and lower bounds of the β_j s. This is comparable to the multiplier bounds imposed for assurance region analysis in Data

Envelopment Analysis (Tompson et al., 1994). Now the quadratic programming model (3.7) can be written as:

$$\begin{aligned}
\min \quad & \frac{1}{2}c\sum_{i=1}^{n} v_i^2 + \frac{1}{2}(1-c)\sum_{i=1}^{n}\left(u_i - \frac{1}{n}\sum_{l=1}^{n} u_l\right)^2 \\
\text{st} \quad & \beta_0 + \sum_{j=1}^{K}\beta_k x_{ij} + v_i - u_i = y_i;\ i = 1,2,\dots,n \\
& s_j^{\min} \le \beta_j \le s_j^{\max};\ \mathrm{j} = 1,2,\dots,\mathrm{k} \\
& \sum_{i=1}^{n} v_i = 0 \\
& u_i \ge 0;\ i = 1,2,\dots,n.
\end{aligned} \qquad (3.9)$$

The optimal solution of this problem yields estimates of the parameters of the specified function. However, these would be merely point estimates subject to variability across samples. It is important, therefore, to examine the sampling distribution of the quadratic problem estimators. Without any specific distributional assumption about the error components and, additionally, due to inequality restrictions, it is not possible to analytically derive the sampling distribution of the estimators and construct confidence intervals. To overcome this difficulty we will bootstrap to form the empirical distribution of the error terms.

3.3 The Bootstrap Procedure

The starting point of any bootstrap procedure is a sample of observed data drawn randomly from some population. The sample statistic computed from this set of observed values is merely an estimate of the corresponding population parameter. When it is not possible to analytically derive the sampling distribution of that statistic, one examines its empirical density function. Unfortunately, however, the researcher has access to only one sample and cannot draw multiple samples from the same underlying population. The basic assumption behind the bootstrap method is that the random sample actually drawn "mimics" its parent population. Therefore, if one draws a random sample with replacement from the observed values in the original sample, it can be treated like a sample drawn from the underlying population itself. Thus, repeated samples with replacement yield different values of the sample statistic under investigation and the associated empirical distribution (over these samples) can provide the sampling distribution of this statistic. This is what is known as the naïve bootstrap.

One major drawback of the naïve bootstrap procedure is that even when sampling with replacement a bootstrap sample will not include *any observation from the parent population, which was not drawn in the initial sample.* Thus, the naïve bootstrap samples are effectively drawn from a discrete population and they fail to reflect the fact that the underlying population density function is continuous. Hence, the empirical distribution derived from the bootstrap sample is not a consistent estimator of the true underlying sampling distribution of the statistic. The use of smoothed bootstrap that it is presented in section (1.5.1) helps to overcome this problem.

For the problem at hand, let F_u and F_v be the unknown population density functions of u and v respectively. The following algorithm describes the steps for the complete method:

i) Solve the mathematical program (3.9) to obtain the estimates

$\hat{\beta}=\{\hat{\beta}_0,\hat{\beta}_1,....,\hat{\beta}_K\}$, $\hat{v}=\{\hat{v}_1,\hat{v}_2,....,\hat{v}_n\}$ and $\hat{u}=\{\hat{u}_1,\hat{u}_2,....,\hat{u}_n\}$.

ii) Select the b-th ($b=1,2,...,B$) independent bootstrap samples

$v_b^* = \{v_{1,b}^*, v_{2,b}^*, ..., v_{n,b}^*\}$ and $u_b^* = \{u_{1,b}^*, u_{2,b}^*, ..., u_{n,b}^*\}$

with replacement from the estimated $\hat{v} = \{\hat{v}_1, \hat{v}_2, ..., \hat{v}_n\}$ and $\hat{u} = \{\hat{u}_1, \hat{u}_2, ..., \hat{u}_n\}$ respectively.

In the smooth bootstrap procedure, we use Gaussian Kernel density functions to smooth the empirical distributions of the error terms. The empirical density function $\hat{f}_v$ of the two-sided error term v_i ($i=1,2,...,n$) can be written as

$$\hat{f}_v = \frac{1}{nh_v}\sum_{i=1}^{n}\phi\left(\frac{t-\hat{v}_i}{h_v}\right) \qquad (3.10)$$

where $\phi(.)$ is the standard density function and h_v is the window width or smoothing parameter for the density function v. The empirical density function $\hat{f}_u$ of the one-sided error term term u_i ($i=1,2,...,n$) is zero for negative u_i s. With the use of the reflection method described in section 1.4.2, the empirical density $\hat{f}_u$ can be written as

$$\hat{f}_u(t) = \frac{1}{2nh_u}\sum_{i=1}^{n}\left[\phi\left(\frac{t-\hat{u}_i}{h_u}\right) - \phi\left(\frac{t+\hat{u}_i}{h_u}\right)\right] \qquad (3.11)$$

where h_u is the window width or smoothing parameter for the density functions u.

The choice of the smoothing parameters h_v and h_u is subjective. It has been shown by Silverman (1986) that the values of h_v and h_u are defined by the minimization of the mean integrated error of the distributions $\hat{f}_v$ and $\hat{f}_u$ respectively. The smoothing parameters are calculated as:

$h_v = 0.9\, A_v\, n^{-1/5}$,

where A_v=min (standard deviation of $\hat{v}$, interquartile range of $\hat{v}/1.34$), (3.12)

and

$h_u = 0.9\, A_u\, n^{-1/5}$,

where A_u=min (standard deviation of $\hat{u}$, interquartile range of $\hat{u}/1.34$). (3.13)

Finally, if we apply the convolution theorem, the smoothed b-th bootstrap samples

$v_b^{**} = \{v_{1,b}^{**}, v_{2,b}^{**}, \ldots, v_{n,b}^{**}\}$ and $u_b^{**} = \{u_{1,b}^{**}, u_{2,b}^{**}, \ldots, u_{n,b}^{**}\}$

can be constructed from $\hat{f}_v$ and $\hat{f}_u$ as following:

$$v_{i,b}^{**} = v_{i,b}^{*} + h_v \varepsilon_{i,b}^{v} \sim \hat{F}_v;\ i = 1,2,\ldots,n \quad (3.14)$$

and

$$u_{i,b}^{**} = \text{absolute}(u_{i,b}^{*} + h_u \varepsilon_{i,b}^{u}) \sim \hat{F}_u;\ i = 1,2,\ldots,n\,. \quad (3.15)$$

where $\varepsilon_{i,b}^{v}$ and $\varepsilon_{i,b}^{u}$ are random numbers, which are identically and independently generated from the standard normal distribution. These random numbers are independent of each other because we assumed that the two error terms are independently distributed.

iii) Create the b-th pseudo data set $(x^i, y_{i,b}^*)$ i=1,2,…,n, where $y_{i,b}^* = f(x^i; \hat{\beta}) + v_{i,b}^{**} - u_{i,b}^{**}$.

iv) Solve model (3.8) using the b-th pseudo-data set to obtain the bootstrap estimates:

$\hat{\beta}_b^* = \{\hat{\beta}_{0,b}^*, \hat{\beta}_{1,b}^*, \ldots, \hat{\beta}_{K,b}^*\}$, $\hat{v}_b^* = \{\hat{v}_{1,b}^*, \hat{v}_{2,b}^*, \ldots, \hat{v}_{n,b}^*\}$ and $\hat{u}_b^* = \{\hat{u}_{1,b}^*, \hat{u}_{2,b}^*, \ldots, \hat{u}_{n,b}^*\}$.

v) Repeat steps (ii)-(iv) B times for a set of estimates $\{\hat{\beta}_b^*, \hat{v}_b^*, \text{and } \hat{u}_b^*$ for $b=1,2,\ldots,B\}$.

vi) Calculate the average of the bootstrap estimates of β, v and u as the arithmetic mean:

$$\bar{\hat{\beta}}_k^{\,*} = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_{k,b}^{*};\ \mathrm{k} = 0,1,\ldots,\mathrm{K}, \tag{3.16a}$$

$$\bar{\hat{v}}_i^{\,*} = \frac{1}{B}\sum_{b=1}^{B}\hat{v}_{i,b}^{*};\ \mathrm{i} = 1,2,\ldots,\mathrm{n}, \tag{3.16b}$$

$$\bar{\hat{u}}_i^{\,*} = \frac{1}{B}\sum_{b=1}^{B}\hat{u}_{i,b}^{*};\ \mathrm{i} = 1,2,\ldots,\mathrm{n}. \tag{3.16c}$$

We can now calculate the bias, bias-corrected estimates and construct confidence intervals for each of the estimated parameters. The estimated bias of the bootstrap estimated parameters based on B replications is

$$\mathrm{bias}_B(\beta_k) = \bar{\hat{\beta}}_k^{*} - \hat{\beta}_k;\ \mathrm{k} = 0,1,2,\ldots,\mathrm{K}, \tag{3.17a}$$

$$\mathrm{bias}_B(v_i) = \bar{\hat{v}}_i^{*} - \hat{v}_i;\ \mathrm{i} = 1,2,\ldots,\mathrm{n}, \tag{3.17b}$$

$$\mathrm{bias}_B(u_i) = \bar{\hat{u}}_i^{*} - \hat{u}_i;\ \mathrm{i} = 1,2,\ldots,\mathrm{n}. \tag{3.17c}$$

Before we correct our estimates for the bias from the bootstrap we need to examine whether the estimated bootstrap-bias is small compared to the estimated standard error. The standard error of each of the estimated parameters is measured as

$$\mathrm{se}_{\mathrm{B}}(\hat{\beta}_k^{*}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\beta}_{k,b}^{*} - \bar{\hat{\beta}}_k^{*})^2};\ \mathrm{k} = 0,1,\ldots,\mathrm{K}, \tag{3.18a}$$

$$\mathrm{se}_{\mathrm{B}}(\hat{v}_{i.}^{*}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{v}_{i,b}^{*} - \bar{\hat{v}}_i^{*})^2};\ \mathrm{i} = 1,2,\ldots,\mathrm{n}, \tag{3.18b}$$

$$\mathrm{se_B}(\hat{u}_i^*) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{u}_{i,b}^* - \bar{\hat{u}}_i^*)^2}\,;\ \mathrm{i} = 1,2,\ldots,\mathrm{n}. \qquad (3.18c)$$

The bias-corrected estimated parameter will be given by the formulas:

$$\hat{\hat{\beta}}_j = \hat{\beta}_j - \mathrm{bias_B}(\beta_j) = 2\hat{\beta}_j - \bar{\hat{\beta}}_j^*;\ \mathrm{j} = 0,1,\ldots,\mathrm{k}, \qquad (3.19a)$$

$$\hat{\hat{v}}_i = \hat{v}_i - \mathrm{bias_B}(v_i) = 2\hat{v}_i - \bar{\hat{v}}_i^*;\ \mathrm{i} = 1,2,\ldots,\mathrm{n}, \qquad (3.19b)$$

$$\hat{\hat{u}}_i = \hat{u}_i - \mathrm{bias_B}(u_i) = 2\hat{u}_i - \bar{\hat{u}}_i^*;\ \mathrm{i} = 1,2,\ldots,\mathrm{n}. \qquad (3.19c)$$

We are also interested in the distribution of the error terms. Because the $\hat{v}$ s by construction add up to 0 in each replication, then the bias of the $\hat{v}$ s will also add up to 0. Thus, the average of the bias-corrected $\hat{v}$ s will be equal to 0. The detailed calculations are presented in Appendix 3.1. From (3.19b) we can calculate the variance of the two-sided bias-corrected error term:

$$\mathrm{var}(\hat{\hat{v}}) = \frac{1}{n-1}\sum_{i=1}^{n}\hat{\hat{v}}_i^2. \qquad (3.20)$$

Similarly we can use (3.18c) to estimate the arithmetic mean and variance of the technical inefficiency error term:

$$\bar{\hat{\hat{u}}} = \frac{1}{n}\sum_{i=1}^{n}\hat{\hat{u}}_i\text{, and} \qquad (3.21a)$$

$$\mathrm{var}(\hat{\hat{u}}) = \frac{1}{n-1}\sum_{i=1}^{n}(\hat{\hat{u}}_i - \bar{\hat{\hat{u}}})^2. \qquad (3.21b)$$

Finally, we can construct confidence intervals for the estimates of the parameter β. First, we must adjust the estimates of β from each bootstrap $\hat{\hat{\beta}}_b^*$, $(b = 1,2,\ldots,B)$ such that they are centered on $\hat{\hat{\beta}}$, the bias-corrected estimate of β, i.e. $E(\hat{\hat{\beta}}_b^*) = \hat{\hat{\beta}}, (b=1,2,\ldots,B)$. According to this, the adjusted estimate from each bootstrap will be

$$\hat{\hat{\beta}}^*_{j,b} = \hat{\beta}^*_{j,b} - 2\,\text{bias}_B(\hat{\beta}_j);\ b=1,2,\ldots,B;\ j=0,1,\ldots,k. \tag{3.22}$$

Once we have calculated the adjusted estimates we can use the percentile method to construct the (1-2a)% confidence intervals for each β as

$$(\hat{\hat{\beta}}_j^{*(a)}, \hat{\hat{\beta}}_j^{*(1-a)});\ j=0,1,\ldots,k, \tag{3.23}$$

where $\hat{\hat{\beta}}_j^{*(a)}$ and $\hat{\hat{\beta}}_j^{*(1-a)}$ is the $(100*a^{th})$ and $(100*(1-a)^{th})$ percentiles of the empirical density of $\hat{\hat{\beta}}^*_{j,b}$ $(b=1,2,\ldots,B;\ j=0,1,\ldots,k)$.

3.4 The Empirical Application

In this chapter we estimate a Cobb Douglas stochastic frontier production function for U.S. manufacturing using the Quadratic Programming-bootstrap procedure described in the previous section. The manufacturing sector is visualized as an industry producing a scalar output from six inputs: (1) production workers (L), (2) non-production workers (employees, EM)), (3) buildings and structures (i.e., land and buildings, BS), (4) machinery and equipment (ME), (5) materials (M), and (6) energy (ENER). The data for different states have been obtained from the 1992 Census of Manufacturing. Details of data construction are provided in Table 3.1. The number of establishments covered by the Census varies widely across states. State level input-output quantity data for the "representative establishment" were constructed by dividing the state-level total values of the variables by the number of establishments covered in the state.

In order to estimate a Cobb-Douglas production function we regressed the logarithm of the output on the logarithms of the inputs and the results are presented in Table 3.2. We find that the coefficients of production workers (L) and energy (ENER) are negative and non-significant. One might want to impose restrictions on these coefficients, such that the estimated values are consistent with the economic theory, for example, non-negative shares for the Cobb-Douglas production function. However, when inequality restrictions are imposed the statistical distributions of the estimated parameters are not well defined (Yancey, 1981, Judge et al., 1985).

We calculated the observed minimum and maximum input shares across all states and the results for each input are presented in Table 3.3. The calculation of the shares was possible due to the availability of input prices in our sample. We notice that the estimated shares of production workers (L), and energy (ENER) are lower than the minimum observed across all states. Similarly, the estimated shares of building and structure (BS) and machinery and equipment (ME) are greater than the observed maximum shares across all states.

We allow the values of the weight parameter, c, in the objective function to vary and we set it equal to 0.1, 0.25, 0.50, 0.75, and 0.90. The optimal solution of (3.9) for the alternative values of c gives us the point estimates, which are presented in Table 3.4. The parameters of interest include the intercept, β_0, the input coefficients, β_L, β_{EM}, β_{BS}, β_{ME}, β_M, β_{ENER}, the variance of the two-sided error term, var(v), the arithmetic mean and the variance of the one-sided error term, $\bar{u}$ and var(u). Due to the log-linearity of our production function, the coefficients of all inputs remain the same across alternative values of the weight parameter. The coefficients of production workers (L) and Energy (ENER) are restricted by their lower limits, i.e. their minimum observed shares, while the coefficients of non-production workers (EM), buildings and structures (BS), and machinery and equipment (ME) are restricted by their upper limits, i.e. their maximum observed shares. Only the values of the intercept are changing. More specifically, the intercept is increasing as the weight, c, is shifted from the technical efficiency ($c<0.5$) to the random shock ($c>0.5$). Also, we observe a pattern for the average and variance of the error terms. The values of the mean and the variance of the one-sided error term that represent the technical inefficiency are increased as the value of the weight parameter, c, is increased, while the variance of the two-sided error tem that represents the random shock is reduced.

Table 3.4 also includes the results for the deterministic frontier ($v=0$) and the case where we ignore the existence of the technical inefficiency ($u=0$). The last case is equivalent to the restricted OLS estimation. As one should expect, the coefficient of the deterministic frontier is higher than the coefficient of the restricted OLS function. This finding agrees with the transformation of a deterministic frontier, where we subtract the arithmetic mean of the one-sided error term from the intercept and the error terms, such that the transformed error terms have zero mean.

At the bottom of Table 3.4, we report the values of the smoothing parameters, h_v and h_u, which are later used for the bootstraps. The values of the smoothing parameters are calculated using (3.12) and (3.13) for each case that we estimate. For each of the estimated scenarios we performed 2000 bootstrap replications based on the algorithm, which is given in the previous section. Tables 3.5 to 3.11 contain for the two alternative scenarios: i) the point estimates for coefficients of the Cobb-Douglas production function from the quadratic program, ii) average estimates from the 2,000 bootstrap estimations (see 3.16), iii) the bias of the point estimates from the average bootstrap estimates (see 3.17), iv) the bias-corrected estimates (see 3.19), and finally v) the 95% confidence interval for the bias-corrected estimates (see 3.23). The bias corrected estimates for non-production workers (EM), building and structures (BS), machinery and equipment (ME), and materials (M) are higher than the point estimates due to negative estimated bias, while the bias corrected estimates for production workers (L) and energy (ENER) are lower than the point estimates due to the existence of positive bias. The variance of the two-sided error terms (v) and the average and variance of the one-sided error term (u) are calculated from the point estimates for the error terms, while the bias corrected variances and arithmetic mean are calculated from bias corrected error terms.

Consider the case of c=0.25. This implies that greater emphasis is given to the technical inefficiency term (u) rather than the random shock (v). The two-sided error term values (v) are allowed to vary more than the one-sided error term values (v), since the influence of the v s on the objective function is diminished. It is interesting to compare the estimated input shares from the regression with the 95% confidence intervals obtained from the bootstrap replications and the adjustment for the bias. With the unique exception of materials (M), no other regression coefficient is within the limits of this confidence interval. On the other hand, all the bias corrected input coefficients are within the 95% confidence interval obtained from the regression. The same conclusions hold true for the other scenarios with alternative values of the weight parameter. Also, we notice that the 95% confidence intervals for all inputs have the smallest width when the weight parameter equals 0.5, and they become wider as the weight parameter approaches 0 or 1.

3.5 Summary

This chapter revives the mathematical programming approach for the estimation of a production function by Aigner and Chu (1968) and develops another method for the estimation of a composed error frontier. The estimation of a composed error frontier with econometric techniques requires the specification of distributional assumptions about the error terms. The principal contribution of this chapter is the estimation of a composed error frontier without the need of such assumptions that are arbitrary and lead to different conclusions when they are altered. Another advantage of this method over the econometric techniques is the ease of imposing inequality restrictions. These are achieved through the use of mathematical programming. However, the resulting solution contains point estimates with no statistical properties. We overcome this problem with the use of bootstrapping. Sections (3.2) and (3.3) describe the quadratic programming and the appropriate bootstrap algorithm. Finally, an empirical application is discussed in section (3.4) as an illustration of this method.

Table 3.1: Definition and construction of the Variables		
Output:	V	= Value of Output (millions of $)
Input 1:	L	= Number of Production Workers (in thousands)
Input 2:	EM	= Number of Non-production Workers (in thousands)
Input 3:	BS	= Building and Structures
		= [Depreciation Expense +
		Rental Payment for Building and Structures]
		/ [Median Rent per Room in 1990] (in millions of $)
Input 4:	ME	= Machinery and Equipment
		= [Depreciation Expense +
		Rental Payment for Machinery and Equipment]
		/ [Price of Machinery and Equipment = $1]
		(in millions of $)
Input 5:	M	= Materials
		= [Material Expenses + Contractual Work]
		/ [Price of Materials = $1] (in millions of $)
Input 6:	ENER	= Energy Input
		= [Expenditure in Fuels and Purchased Electricity]
		/ [Price of Energy per Btu] (in millions of $)

where: Price of Energy = [Share of Coal in Total Expenditure on Energy in the State] *[Price of Coal] + [Share of Natural Gas in Total Expenditure on Energy in the State] *[Price of Natural Gas] + [Share of Petroleum in Total Expenditure on Energy in the State] *[Price of Petroleum] + [Share of Electricity in Total Expenditure on Energy in the State] *[Price of Electricity]
Note: These data are from Ray and Mukherji (1998). All variables refer to the year 1992 unless otherwise stated. The Census of Manufacturers reports the input-output data aggregated for all establishments covered from a state in the Census. The input-output data are scaled by the reported number of establishments for each state.

Table 3.2: Regression Analysis

Model: MODEL1
Dependent Variable: V

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Prob>F
Model	6	8.20377	1.36729	268.218	0.0001
Error	44	0.22430	0.00510		
C Total	50	8.42807			

Root MSE	0.07140	R-square	0.9734
Dep Mean	1.96180	Adj R-sq	0.9698
C.V.	3.63942		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	T for H0: Parameter=0	Prob > |T|
INTERCEP	1	2.955038	0.39346284	7.510	0.0001
L	1	-0.036681	0.05118087	-0.717	0.4774
EM	1	0.229876	0.04125152	5.573	0.0001
BS	1	0.103130	0.04784336	2.156	0.0366
ME	1	0.157231	0.05268956	2.984	0.0046
M	1	0.599548	0.04680768	12.809	0.0001
ENER	1	-0.031431	0.02943688	-1.068	0.2915

	95% Confidence	Interval
INTERCEP	2.18387	3.72621
L	-0.13699	0.06363
EM	0.14902	0.31073
BS	0.00936	0.19690
ME	0.05396	0.26050
M	0.50781	0.69129
ENER	-0.08913	0.02626

Table 3.3: Minimum and maximum input shares

Variables	minimum share	maximum share
L	0.05427	0.13600
EM	0.05708	0.26876
BS	0.00399	0.01783
ME	0.01340	0.05267
M	0.21086	0.61335
ENER	0.00452	0.04562

Table 3.4:Quadratic Programming Results (point estimates)

	c = 0.10	c = 0.25	c = 0.50	c = 0.75	c = 0.90	v=0	u=0
β_0=intercept	2.84466	2.89146	2.96945	3.04745	3.09425	3.12643	2.81346
$\beta_1 = \beta_L$	0.05427	0.05427	0.05427	0.05427	0.05427	0.05427	0.05427
$\beta_2 = \beta_{EM}$	0.26849	0.26849	0.26849	0.26849	0.26849	0.26849	0.26849
$\beta_3 = \beta_{BS}$	0.01783	0.01783	0.01783	0.01783	0.01783	0.01783	0.01783
$\beta_4 = \beta_{ME}$	0.05267	0.05267	0.05267	0.05267	0.05267	0.05267	0.05267
$\beta_5 = \beta_M$	0.58908	0.58908	0.58908	0.58908	0.58908	0.58908	0.58908
$\beta_6 = \beta_{ENER}$	0.00452	0.00452	0.00452	0.00452	0.00452	0.00452	0.00452
Variance(v)	0.00427	0.00297	0.00132	0.00033	0.00005	N.A.	0.00528
Average (u)	0.03120	0.07800	0.15599	0.23399	0.28078	0.31296	N.A.
Variance(u)	0.00005	0.00033	0.00132	0.00297	0.00427	0.00528	N.A.
h_v	0.01992	0.01660	0.01107	0.00553	0.00221	N.A.	
h_u	0.00221	0.00553	0.01107	0.01660	0.01992		N.A.

Table 3.5: Bootstrap Results for c = 0.10						
	Point Estimates	Average Bootstrap Estimate	Bias	Bias Corrected Estimates	95% confidence interval for the Bias Corrected Estimates	
Intercept	2.84466	2.86589	0.02123	2.82343	2.55938	3.15153
β_L	0.05427	0.07735	0.02308	0.03120	0.00812	0.08985
β_{EM}	0.26849	0.26118	-0.00731	0.27580	0.23658	0.28338
β_{BS}	0.01783	0.01144	-0.00640	0.02423	0.01678	0.03062
β_{ME}	0.05267	0.03992	-0.01275	0.06542	0.03891	0.07817
β_M	0.58908	0.57695	-0.01214	0.60122	0.53624	0.63762
β_{ENER}	0.00452	0.01206	0.00755	-0.00303	-0.01058	0.03053
Variance(v)	0.00428			0.01676		
Average (u)	0.03120			0.04050		
Variance(u)	0.00005			0.00021		

Table 3.6: Bootstrap Results for c = 0.25

	Point Estimates	Average Bootstrap Estimate	Bias	Bias Corrected Estimates	95% confidence interval for the Bias Corrected Estimates	
Intercept	2.89146	2.87653	-0.01493	2.90639	2.68315	3.19200
β_L	0.05427	0.07574	0.02146	0.03281	0.01135	0.08915
β_{EM}	0.26849	0.26199	-0.00650	0.27499	0.24120	0.28176
β_{BS}	0.01783	0.01114	-0.00670	0.02453	0.01738	0.03123
β_{ME}	0.05267	0.03929	-0.01338	0.06604	0.04015	0.07942
β_M	0.58908	0.57994	-0.00915	0.59823	0.54230	0.63164
β_{ENER}	0.00452	0.01098	0.00647	-0.00195	-0.00842	0.02535
Variance(v)	0.00297			0.01171		
Average (u)	0.07800			0.10930		
Variance(u)	0.00033			0.00130		

Table 3.7: Bootstrap Results for c = 0.50

	Point Estimates	Average Bootstrap Estimate	Bias	Bias Corrected Estimates	95% confidence interval for the Bias Corrected Estimates	
Intercept	2.96945	2.90000	-0.06945	3.03891	2.82565	3.30702
β_L	0.05427	0.07382	0.01955	0.03472	0.01518	0.08340
β_{EM}	0.26849	0.26251	-0.00599	0.27448	0.24319	0.28073
β_{BS}	0.01783	0.01126	-0.00657	0.02440	0.01713	0.03097
β_{ME}	0.05267	0.03987	-0.01280	0.06547	0.03900	0.07826
β_M	0.58908	0.58155	-0.00753	0.59661	0.54385	0.62841
β_{ENER}	0.00452	0.01060	0.00609	-0.00157	-0.00766	0.02273
Variance(v)	0.00132			0.00519		
Average (u)	0.15599			0.23557		
Variance(u)	0.00132			0.00519		

Table 3.8: Bootstrap Results for c = 0.75						
	Point Estimates	Average Bootstrap Estimate	Bias	Bias Corrected Estimates	95% confidence interval for the Bias Corrected Estimates	
Intercept	3.04745	2.96463	-0.08282	3.13026	2.88197	3.43647
β_L	0.05427	0.07470	0.02042	0.03385	0.01342	0.08971
β_{EM}	0.26849	0.26285	-0.00564	0.27413	0.23940	0.28003
β_{BS}	0.01783	0.01123	-0.00660	0.02444	0.01720	0.03104
β_{ME}	0.05267	0.03940	-0.01327	0.06593	0.03993	0.07920
β_M	0.58908	0.58049	-0.00859	0.59767	0.54219	0.63052
β_{ENER}	0.00452	0.01127	0.00676	-0.00224	-0.00900	0.02453
Variance(v)	0.00033			0.00130		
Average (u)	0.23399			0.33420		
Variance(u)	0.00297			0.01169		

Table 3.9: Bootstrap Results for c = 0.90

	Point Estimates	Average Bootstrap Estimate	Bias	Bias Corrected Estimates	95% confidence interval for the Bias Corrected Estimates	
Intercept	3.09425	3.02093	-0.07332	3.16756	2.87585	3.51009
β_L	0.05427	0.07701	0.02273	0.03154	0.00881	0.09053
β_{EM}	0.26849	0.26096	-0.00753	0.27602	0.23620	0.28381
β_{BS}	0.01783	0.01141	-0.00643	0.02426	0.01684	0.03069
β_{ME}	0.05267	0.03889	-0.01378	0.06645	0.04097	0.08024
β_M	0.58908	0.57952	-0.00956	0.59864	0.53779	0.63247
β_{ENER}	0.00452	0.01161	0.00709	-0.00258	-0.00967	0.02674
Variance(v)	0.00005			0.00021		
Average (u)	0.28078			0.37481		
Variance(u)	0.00428			0.01681		

Table 3.10: Bootstrap Results for v_i=0 for every i=1,2,...,n

	Point Estimates	Average Bootstrap Estimate	Bias	Bias Corrected Estimates	95% confidence interval for the Bias Corrected Estimates	
Intercept	3.12643	1.21007	-1.91636	5.04278	3.83271	7.23022
β_L	0.05427	0.03119	-0.02308	0.07735	0.04616	0.18082
β_{EM}	0.26849	0.10251	-0.16598	0.43447	0.33196	0.60072
β_{BS}	0.01783	0.00467	-0.01316	0.03100	0.02633	0.04416
β_{ME}	0.05267	0.01503	-0.03764	0.09031	0.07529	0.12796
β_M	0.58908	0.22561	-0.36347	0.95255	0.72694	1.34029
β_{ENER}	0.00452	0.00467	0.00016	0.00436	-0.00031	0.03480
Average (u)	0.31296			0.53852		
Variance(u)	0.00528			0.02073		

Table 3.11: Bootstrap Results for u_i=0 for every i=1,2,...,n

	Point Estimates	Average Bootstrap Estimate	Bias	Bias Corrected Estimates	95% confidence interval for the Bias Corrected Estimates	
Intercept	2.813462	2.785478	0.027984	2.785478	2.501476	3.132848
β_L	0.054274	0.029511	0.024763	0.029511	0.004748	0.086473
β_{EM}	0.268491	0.27704	-0.00855	0.27704	0.231737	0.285858
β_{BS}	0.017834	0.024376	-0.00654	0.024376	0.017074	0.030918
β_{ME}	0.05267	0.067106	-0.01444	0.067106	0.042277	0.081543
β_M	0.589083	0.600621	-0.01154	0.600621	0.534427	0.636425
β_{ENER}	0.004515	-0.00353	0.008043	-0.00353	-0.01157	0.029535
Variance(v)	0.00528			0.02081		

Table 3.12: Data

obs	v	l	em	bs	me	mc	ener
1	8.2572	0.044045	0.014848	0.001427	0.25796	4.1684	0.03118
2	7.1181	0.023669	0.007101	0.000506	0.20237	4.083	0.03286
3	5.3844	0.021087	0.016471	0.000541	0.15661	2.1395	0.00718
4	8.7708	0.045719	0.01219	0.000966	0.22014	4.7153	0.0201
5	5.9327	0.022092	0.016479	0.000515	0.16738	2.5902	0.00686
6	5.5128	0.01977	0.014464	0.000579	0.13482	2.4787	0.008
7	6.3889	0.027221	0.023846	0.000538	0.17822	2.2152	0.00635
8	17.7167	0.042334	0.048168	0.000749	0.35834	10.4651	0.02789
9	4.4072	0.008297	0.020087	0.000493	0.09563	0.9293	0.00121
10	3.9262	0.017605	0.011232	0.000393	0.10792	1.7552	0.0072
11	9.2876	0.04034	0.016513	0.000728	0.2317	4.7781	0.01976
12	3.7374	0.012647	0.007549	0.000371	0.07814	1.9261	0.00284
13	5.8745	0.024932	0.011184	0.000885	0.1802	3.169	0.02067
14	8.4058	0.031226	0.020374	0.000734	0.21677	4.0254	0.01463
15	11.3526	0.04681	0.020047	0.001285	0.33154	5.503	0.03144
16	11.815	0.040276	0.017812	0.001051	0.24163	6.0035	0.02718
17	10.4075	0.036487	0.017825	0.001086	0.2094	5.6884	0.01876
18	13.8676	0.046924	0.017181	0.001041	0.28718	7.0584	0.04122
19	15.1141	0.031077	0.013068	0.001136	0.43197	9.2702	0.10789

20	5.3115	0.030318	0.011091	0.001111	0.21364	2.3552	0.01828
21	7.15	0.026391	0.018379	0.000629	0.19245	3.1956	0.0151
22	6.4034	0.026999	0.020363	0.000856	0.17705	2.531	0.00579
23	9.6275	0.034741	0.020719	0.000879	0.24166	5.0341	0.0135
24	7.2206	0.028344	0.020994	0.000596	0.18535	3.4894	0.01097
25	8.7361	0.049867	0.013417	0.001127	0.23175	4.5963	0.02217
26	9.2734	0.033049	0.0191	0.000746	0.15388	4.72	0.01333
27	3.019	0.01141	0.004288	0.00035	0.06526	1.7887	0.02228
28	10.7881	0.035422	0.013962	0.000711	0.1446	6.2015	0.01554
29	2.689	0.014171	0.007846	0.000319	0.10184	1.1323	0.00602
30	4.8364	0.025912	0.0142	0.00053	0.1689	1.8446	0.00481
31	6.5414	0.022855	0.020534	0.000515	0.15037	2.7226	0.00981
32	5.3512	0.016928	0.007712	0.000842	0.1094	2.7275	0.0105
33	5.7122	0.02223	0.017085	0.000636	0.16363	2.2301	0.00731
34	10.8546	0.05131	0.018716	0.001019	0.25026	4.9404	0.01831
35	5.2859	0.018891	0.008696	0.000518	0.14123	3.1663	0.01635
36	10.0225	0.037233	0.019894	0.001045	0.2308	4.8511	0.02604
37	7.4102	0.026747	0.011614	0.000708	0.16284	3.7284	0.02129
38	4.7141	0.02104	0.009835	0.000504	0.13538	2.3499	0.01536
39	7.6928	0.033601	0.018907	0.000759	0.18576	3.4622	0.01474
40	3.5574	0.022131	0.011065	0.000423	0.09209	1.4362	0.00571
41	10.8181	0.056375	0.019446	0.001157	0.34648	5.0809	0.03657

42	6.78	0.028459	0.011136	0.000994	0.1099	4.0206	0.00822
43	10.0773	0.048212	0.01739	0.000985	0.28077	4.8329	0.01991
44	9.822	0.026881	0.017154	0.000801	0.28424	5.3972	0.04413
45	6.1665	0.026337	0.014772	0.000836	0.16238	2.9833	0.0163
46	4.7367	0.022057	0.011103	0.00046	0.24948	1.8513	0.00598
47	10.1611	0.04347	0.018945	0.000712	0.23755	4.2401	0.02136
48	8.5364	0.023504	0.016322	0.000744	0.16075	4.9492	0.0301
49	7.4723	0.031352	0.012675	0.000843	0.22872	3.3199	0.04817
50	8.7849	0.036621	0.017508	0.000976	0.20659	4.2414	0.01719
51	4.1237	0.011073	0.004498	0.000366	0.16799	2.4042	0.02559
average	7.70497	0.02993	0.01560	0.00076	0.19649	3.81934	0.01961
standard deviation	3.11636	0.01147	0.00659	0.00027	0.07460	1.89792	0.01666

APPENDIX 3.1

From the solution of (3.4) we get for firm i:

$$y_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i,1} + \ldots + \hat{\beta}_k x_{i,k} + \hat{v}_i - \hat{u}_i \ . \qquad (3.24)$$

From b-th bootstrap we get for firm i:

$$y_i = \hat{\beta}_{0,b} + \hat{\beta}_{1,b} x_{i,1} + \ldots + \hat{\beta}_{k,b} x_{i,k} + \hat{v}_{i,b} - \hat{u}_{i,b} \qquad (3.25)$$

The average over all the bootstrap replications is

$$y_i = \bar{\hat{\beta}}_{0,b} \cdot + \bar{\hat{\beta}}_{1,b} x_{i,1} + \ldots + \bar{\hat{\beta}}_{k,b} x_{i,k} + \bar{\hat{v}}_{i,b} - \bar{\hat{u}}_{i,b} \qquad (3.26)$$

If we subtract (3.23) from (3.25) we get

$$0 = (\bar{\hat{\beta}}_{0,b} - \hat{\beta}_0) \cdot + (\bar{\hat{\beta}}_{1,b} - \hat{\beta}_1) x_{i,1} + \ldots + (\bar{\hat{\beta}}_{k,b} - \hat{\beta}_k) x_{i,k} + (\bar{\hat{v}}_{i,b} - \hat{v}_i) - (\bar{\hat{u}}_{i,b} - \hat{u}_i) \ (3.27)$$

which can be re-written as

$$0 = \text{bias}_B(\beta_0) + \text{bias}_B(\beta_1) x_{i,1} + \ldots + \text{bias}_B(\beta_k) x_{i,k} + \text{bias}_B(v_i) - \text{bias}_B(u_i) \qquad (3.28)$$

If we subtract two times (3.27) from (3.25) we get an expression for the bias corrected estimates.

$$y_i = (\bar{\hat{\beta}}_{0,b} - \hat{\beta}_{0,bias}) \cdot + (\bar{\hat{\beta}}_{1,b} - \hat{\beta}_{1,bias}) x_{i,1} + \ldots + (\bar{\hat{\beta}}_{k,b} - \hat{\beta}_{k,bias}) x_{i,k} + (\bar{\hat{v}}_{i,b} - \hat{v}_{i,bias}) - (\bar{\hat{u}}_{i,b} - \hat{u}_{i,bias})$$
(3.29)

or equivalently

$$y_i = \hat{\hat{\beta}}_{0,b} + \hat{\hat{\beta}}_{1,b} x_{i,1} + \ldots + \hat{\hat{\beta}}_{k,b} x_{i,k} + \hat{\hat{v}}_{i,b} - \hat{\hat{u}}_{i,b} \qquad (3.30)$$

The average of the above expression over all the bootstraps for firm i becomes:

$$y_i = \bar{\hat{\hat{\beta}}}_0 + \bar{\hat{\hat{\beta}}}_1 x_{i,1} + \ldots + \bar{\hat{\hat{\beta}}}_k x_{i,k} + \bar{\hat{\hat{v}}}_i - \bar{\hat{\hat{u}}}_i \qquad (3.31).$$

CHAPTER 4: A BOOTSTRAP PROCEDURE TO CAPTURE UNIT SPECIFIC EFFECTS IN DATA ENVELOPMENT ANALYSIS

4.1 Introduction

One major drawback of Data Envelopment Analysis (DEA) is that it is non-statistical and the efficiency score obtained for an individual Decision Making Unit (DMU) is a point estimate without any confidence interval around it. In recent years, researchers have resorted to bootstrapping (e.g. Simar (1992, 1996), Simar and Wilson (1997a, 1997b) among others) in order to generate empirical distributions of efficiency scores from repeated applications of DEA after resampling. The essential procedure is to pool the efficiency measures obtained from the actual data and then randomly sample with replacement from this pool to construct pseudo-data on inputs (or outputs) for the DMUs. These artificial data on inputs (outputs) are associated with actual output (input) data for another round of DEA. Repeating this procedure a large number of times generates large enough samples of efficiency scores for each DMU. Then one can look at the mean and the variance of each of the empirical distributions of efficiency.

While this procedure is quite appealing and is gaining wide acceptance, in a sense, it goes to the other extreme by assuming that all DMUs have the same probability of getting an efficiency score from any specified interval within the [0,1] range. This reduces efficiency to a purely random variable and there would be little point in talking of the efficiency of one DMU relative to the others. In reality, however, some DMUs are more likely to be rated at a higher efficiency level than other DMUs. There usually are systematic factors that contribute to differences in efficiency. The existing bootstrapping procedures do not consider the possibility that the distributions of efficiency conditional on unit specific factors may differ across DMUs. One can argue in favor of including these factors within the scope of the DEA model itself so that the remaining variation in efficiency can be justifiably attributed to purely random factors. However, inclusion of these factors as non-discretionary inputs within the DEA model automatically extends the disposability property (weak or strong) to such variables. This is not a realistic assumption in many situations. This is one reason why researchers often regress DEA efficiency scores on a number of explanatory variables to adjust for environmental factors and they do not include these factors in the DEA model itself (e.g. Ray(1991), McCarty and Yaisawarng (1993)).

The essay in this chapter proposes a bootstrap procedure that empirically generates the conditional distribution of efficiency for each individual DMU. The principal innovation in this study is that instead of resampling directly from the pooled DEA scores, we first regress these scores on a set of explanatory variables not included at the DEA stage and we bootstrap the residuals from this regression. These pseudo-efficiency scores incorporate the systematic effects of unit-specific factors along with the contribution of the randomly drawn residual.

This chapter is organized as follows. In section 4.2 we set up the DEA model and the regression of the technical efficiency on the unit-specific factors. Section 4.3 describes the two-step bootstrap procedure and differentiates it from the one-step bootstrap. Section 4.4

reports the findings from an empirical application using data from Connecticut public schools. Finally, the last section summarizes the developed methodology.

4.2 DEA and Regression models

Consider an industry producing a vector of *m* outputs, $y=(y_1, y_2, ..., y_m)$, from a vector of *k* inputs, $x=(x_1, x_2, ..., x_k)$. Let the vectors x^i and y^i represent, respectively, the input and output bundles of the i-th firm or decision making unit (DMU). Suppose that the input-output data are observed for *n* DMUs. As shown in section (1.4), the output-oriented technical efficiency of the j-th firm under variable returns to scale (VRS) can be computed by solving the linear programming (LP) problem:

$$\begin{aligned} & \max \phi_j \\ s.t. \quad & \sum_{i=1}^{n} \lambda_i y_t^i \geq \phi_j y_t^j; \ \text{for } t = 1,2,...,m \\ & \sum_{i=1}^{n} \lambda_i x_s^i \leq x_s^j; \ \text{for } s = 1,2,...,k \\ & \sum_{i=1}^{n} \lambda_i = 1; \\ & \lambda_i \geq 0 \quad \text{for } i = 1,2,...,n. \end{aligned} \tag{4.1}$$

The technical efficiency for the j-th firm will be the inverse of ϕ_j.

$$\mathrm{TE}_j = \frac{1}{\phi_j} \tag{4.2}$$

When ϕ_j is equal to 1, the technical efficiency is equal to 1, i.e. the firm is 100% efficient. If ϕ_j is greater than 1, then the firm is technically inefficient and the technical efficiency measure is less than 1.

DEA models lead to measures of technical efficiency that are point estimates and therefore lack statistical properties. This problem has been addressed with the use of bootstrap methods. Simar (1992, 1996) and Simar and Wilson (1997a, 1997b) set the foundation of consistent use of bootstrapping to generate empirical distributions of efficiency scores. This method has been described in section 1.6. The problem with bootstrapping the technical efficiency measures is the assumption that all DMUs have the same probability of getting an efficiency score from any specified interval within the (0-1) range. There usually are systematic factors that contribute to differences in efficiency and can lead to different technical efficiency scores. For example, for an inter-country analysis of manufacturing production it is not sensible to conceptualize a data generating process where Germany and Ethiopia have the same probability of getting efficiency scores in excess of 0.975. The existing bootstrapping procedures do not consider the possibility that the distributions of efficiency conditional on unit specific factors may differ across DMUs. One can argue in favor of including these factors within the scope of the DEA model itself so that the remaining variation in efficiency can be justifiably attributed to purely random factors. However, inclusion of these factors as non-discretionary inputs within the DEA model automatically extends the disposability property (weak or strong) to such variables. This is not a realistic assumption in many situations. For example, in the study for Connecticut schools that is later presented in this chapter, proportion of minorities in the student population and parental education are socioeconomic conditions in the home-life of a pupil. These variables can influence the students, but a researcher cannot assume that free disposability is applicable. This is one reason why researchers often regress DEA efficiency scores on a number of explanatory variables to adjust for environmental factors; they do not include these factors in the DEA model itself.

Let the vector z^i represent such characteristics of the i-th firm. A regression permits us to determine the proportion of the technical efficiency that is due to these characteristics and the proportion that is due to random error:

$$\phi_i = \alpha + z^i \gamma + u_i, \tag{4.3}$$

where $(\alpha + z^i\gamma)$ is the component of technical efficiency that varies systematically with the DMU characteristics and u_i is the random shock. If we estimate the above regression with Ordinary Least Squares, some of the estimated $\hat{\phi}(=\hat{\alpha} + z^i\hat{\gamma})$s may be less than 1, or equivalently $(\hat{\phi} - 1) \leq 0$, which violates the natural restriction on efficiency. As we get a consistent estimate of $\hat{\phi}$ and $\hat{\gamma}$, we can adjust the estimated technical efficiency

$$\tilde{\hat{\phi}}_i - 1 = \hat{\phi}_i - \min_j\{\hat{\phi}_j\} \geq 0, \tag{4.4}$$

or, equivalently,

$$\begin{aligned} \tilde{\hat{\phi}}_i - 1 &= \hat{\alpha} + z^i\hat{\gamma} - \min_j\{\hat{\alpha} + z^j\hat{\gamma}\} = \\ &= [z^i - \min_j\{z^j\}]\hat{\gamma} \Rightarrow \\ \tilde{\hat{\phi}}_i &= 1 + [z^i - \min_j\{z^j\}]\hat{\gamma}. \end{aligned} \tag{4.5}$$

Now the adjusted error term, $\tilde{\hat{u}}_i$, will be

$$\begin{aligned}\tilde{\hat{u}}_i &= \phi_i - \tilde{\hat{\phi}}_i \\ &= \phi_i - (\hat{\phi}_i - \min_j\{\hat{\phi}_j\} + 1) = \\ &= \hat{u}_i + \min_j\{\hat{\phi}_j\} - 1.\end{aligned} \tag{4.6}$$

Now bootstrap from the estimated sample of $\tilde{\hat{u}}_i$ s.

4.3 A Two Step Bootstrap Procedure

The bootstrap algorithm that generates the distribution of efficiency for each individual DMU, conditional on unit specific factors, can be described as follows:

i) For each DMU *i* compute ϕ from the DEA model in (4.1), for i=1,2,...,n.

ii) Regress ϕ_i on the DMU characteristics z_i and adjust them using (4.5)

iii) Calculate the residuals $\tilde{\hat{u}}_i = \phi_i - \tilde{\hat{\phi}}_i$ for each i=1,2,...,n (as shown in (4.4) to (4.6)).

iv) Select the b-th ($b=1,2,\ldots,B$) independent bootstrap sample $\hat{u}_b^* = \{\hat{u}_{1b}^*, \hat{u}_{2b}^*, \ldots, \hat{u}_{nb}^*\}$, which consists of n data values drawn with replacement from the observed sample $\tilde{\hat{u}} = \{\tilde{\hat{u}}_1, \tilde{\hat{u}}_2, \ldots, \tilde{\hat{u}}_n\}$.

v) Generate the smoothed bootstrap sample

$$\hat{u}_{ib}^{**} = \hat{u}_{ib}^* + h\varepsilon_i; \quad \varepsilon_i \sim N(0,1) \quad \text{for } i = 1,2,\ldots,n$$

where h is the smoothing parameter. Following Silverman (1986) we can select the value of the window width that minimizes the approximate mean integrated square error, i.e. h is equal to:

$h = 0.9\, A\, n^{-1/5}$,

where $A = min$ (standard deviation of $\hat{u}$, interquartile range of $\tilde{\hat{u}}/1.34$),

vi) Create the b-th pseudo sample (x^i, y_b^{i*}) i=1,2,...,n, where

$y_b^{i*} = (y^i * \phi_i)/\phi_{i,b}^*$ and $\phi_{i,b}^* = \hat{\alpha} + z^i\hat{\gamma} + \hat{u}_{ib}^{**}$ for $i = 1,2,...,n$

vii) Solve the DEA model for each DMU i in (4.1) using the b-th pseudo-data set to obtain new values for ϕ_i, $\hat{\phi}_{i,b}^*$ for i=1,2,...,n.

viii) Repeat steps (iv)-(vii) B-times to obtain the maximum producible output for each DMU i, (i=1,2,...,n) :

$$\hat{y}_{i,b}^f = y_b^{i*}\,\hat{\phi}_{i,b}^*;\ b = 1,2,...,B. \tag{4.7}$$

ix) Calculate the average of the bootstrap estimates of y^f s, the bias and the confidence intervals.

Note that we are constructing confidence intervals for the maximum producible output from the input bundles of each firm through bootstrap and not for the technical efficiency score as is common in the literature. The usual procedure (e.g. Simar and Wilson, 1992, 1995) is to first obtain bootstrap estimates for the ϕ s and then to multiply them by the observed output level, i.e.

$$\hat{y}_{i,b}^{f*} = y^i\,\hat{\phi}_{i,b}^*;\ b = 1,2,...,B;\ i = 1,2,...,n \tag{4.8}$$

The problem with this approach is that it ignores the fact that the bootstrap measure of technical efficiency is calculated with respect the pseudo-output and not the actual output. While the actual level of output remains the same, the pseudo-output level changes in different replications of the pseudo-data. The two methods of calculation generate different empirical distributions of the maximum output producible from the observed input bundle.

The average of the bootstrap estimate for unit *i* is:

$$\bar{y}_i^{\,f} = \frac{1}{B}\sum_{b=1}^{B}\hat{y}_{i,b}^{f};\ i=1,\dots,n. \tag{4.9}$$

We can now calculate the bias, bias-corrected estimates and construct confidence intervals for each of the estimated parameters. The estimated bias of the bootstrap estimated parameters based on *B* replications is

$$\text{bias}_B(\phi_i) = \bar{\hat{\phi}}_i - \phi_i;\ i=1,2,\dots,n,\ \text{and} \tag{4.10a}$$

$$\text{bias}_B(y_i^{\,f}) = \bar{\hat{y}}_i^{\,f} - y_i^{\,f};\ i=1,2,\dots,n. \tag{4.10b}$$

The standard error of each of the estimated technical efficiency and maximum output are measured as

$$\text{se}_B(\phi_i^{*}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\phi}_{k,b}^{*} - \bar{\hat{\phi}}_k^{*})^2};\ i=1,2,\dots,n,\ \text{and} \tag{4.11a}$$

$$\text{se}_B(y_i^{f}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{y}_{i,b}^{f} - \bar{\hat{y}}_i^{f})^2};\ i=1,2,\dots,n. \tag{4.11b}$$

The bias-corrected estimated parameter will be given by the formulas:

$$\hat{\hat{\phi}}_i = \hat{\phi}_i - \text{bias}_B(\hat{\phi}_i) = 2\hat{\phi}_i - \overline{\hat{\phi}}_j^*;\ j = 1,2,\ldots,n, \tag{4.12a}$$

$$\hat{\hat{y}}_i^f = \hat{y}_i^f - \text{bias}_B(\hat{y}_i^f) = 2\hat{y}_i^f - \overline{\hat{y}}_i^f;\ i = 1,2,\ldots,n. \tag{4.12b}$$

Finally, we can construct confidence intervals for the estimates of ϕ and y. First, we must adjust the estimates of ϕ and y from each bootstrap $\hat{\hat{\phi}}_{i,b}^*$, and $\hat{\hat{y}}_{i,b}^f$ $(b = 1,2,\ldots,B)$ such that they are centered around their bias-corrected estimates. Accordingly, the adjusted estimate from each bootstrap will be

$$\hat{\hat{\phi}}_{i,b}^* = \hat{\phi}_{i,b}^* - 2\,\text{bias}_B(\hat{\phi}_i);\ b = 1,2,\ldots,B;\ i = 1,2,\ldots,n, \text{ and} \tag{4.13a}$$

$$\hat{\hat{y}}_{i,b}^f = \hat{y}_{i,b}^f - 2\,\text{bias}_B(\hat{\phi}_i^f);\ b = 1,2,\ldots,B;\ i = 1,2,\ldots,n. \tag{4.13b}$$

Once we have calculated the adjusted estimates we can use the percentile method to construct confidence intervals with length (1-2a)% for the technical efficiency score of each DMU as

$$(\hat{\hat{\phi}}_i^{*(a)}, \hat{\hat{\phi}}_i^{*(1-a)});\ i = 1,2,\ldots,n, \tag{4.14a}$$

where $\hat{\hat{\phi}}_i^{*(a)}$ and $\hat{\hat{\phi}}_i^{*(1-a)}$ is the (100*a^th^) and (100*(1-a)^th^) percentiles of the empirical density of $\hat{\hat{\phi}}_{i,b}^*$ $(b = 1,2,\ldots,B)$ for each $i = 1,2,\ldots,n$. Similarly, the (1-2a)% confidence intervals for the maximum producible output for each DMU as

$$(\hat{\hat{y}}_i^{f(a)}, \hat{\hat{y}}_i^{f(1-a)});\ i = 1,2,\ldots,n. \tag{4.14b}$$

4.4 A Study of Connecticut Public Schools

In our empirical application, data from 118 school districts for the school year 1995-1996 have been used to evaluate efficiency of individual districts. A single output measured by the average score in the Scholastic Aptitude Test (SAT) is used as a measure of output of a school district. This test consists of a verbal and a mathematical part, each with a maximum score of 800 points. In spite of the fact that people have expressed reservations about the appropriateness of the SAT score as an index of cognitive skills acquired by a pupil, in the absence of any nation-wide tests, it has emerged as the most important yardstick of performance by school districts. The school inputs included are (i) regular instruction teachers, (ii) special education teachers, (iii) support staff (like counselors), (iv) administrators, (v) computers, and (vi) total instruction hours during the school year. All personnel inputs are measured as full-time equivalent units per pupil. Computers are also measured as units per pupil.

It is widely agreed that productivity and efficiency of school inputs depend critically on the socioeconomic conditions in the home-life of a pupil. In the present study, we include race and parental education as two significant factors. A higher proportion of minorities in the student population reflects disadvantaged socioeconomic background of the average student and is expected to lower efficiency of the school inputs of the district. On the other hand, a higher proportion of college graduates among the adults in the population district is an indicator of a higher educational level of the parents of the average student and is expected to enhance efficiency. Table 4.1 summarizes the basic statistics of the variables used in this study. The complete data set is in Table 4.9 of this chapter.

In the two-level bootstrap procedure we start with the output oriented DEA model for the variable returns to scale technology using only the school inputs. These DEA scores and the corresponding maximum producible output are shown in Table 4.2. Of the 118 School districts, 23 were rated efficient. On the other hand, large urban districts showed large potential increase in the average SAT scores- by 39% in Bridgeport, 32.51% in Hartford, and 35.68% in New Haven. Given that these are also the districts with the highest proportion of non-whites and also low percentage of college graduates in the districts, such findings are not surprising.

Next we regress these scores on the socioeconomic characteristics of race and parental education:

$$\phi_i = \alpha + \gamma_1 BA_i + \gamma_2 MINO_i + u_i, \tag{4.15}$$

The Ordinary Least Squares results are shown in Table 4.3. As anticipated, higher proportion of minorities (MINO) in the student population leads to higher ϕ s and thus lower efficiency of the school. On the other hand, a higher proportion of college graduates among the adults in the population district (BA) leads to lower ϕ s and thus improves the efficiency of the school.

Before we proceed with the bootstrap procedure, we adjusted the expected ϕ s and the corresponding residuals using (4.4) to (4.6). The bootstrap samples are drawn from these residuals and we apply the bootstrap algorithm presented in the previous section. The bootstrap procedure was repeated 600 times. Table 4.4 reports the results from the bootstrap for the maximum producible output and Table 4.5 reports the results for the inverse of technical efficiency scores. The first column of these tables contains the name of the school district. The columns 2-6 present results on the maximum producible output as it is estimated from the DEA model, the bootstrap average, the standard deviation of the bootstrap average, the bias of the bootstrap average and the bias corrected estimate respectively. The last two columns give the lower and upper limit of the 95% confidence interval for the bias-corrected values. We can compare the actual performance of a school district with the potential one by comparing the bias-corrected maximum producible output, $\hat{\hat{y}}_i^f$, from Table 4.4 with the observed output, y, in Table 4.2. [1]

For comparison, an extended DEA model incorporating the socioeconomic variables was solved for each district to generate an alternative set of ϕ s. The results from the extended DEA model are shown in the last 3 columns of Table 4.2. The two sets of ϕ s obtained from the DEA models have correlation equal to 0.65. This time the DEA with all data resulted in optimal score equal to unity in 54 of the 118 school districts. Bridgeport and Hartford were found to be 100% efficient, while New Haven, which is also a disadvantaged community, has potential increase in the average SAT score by 24.61%, which is lower than previously.

Tables 4.6 and 4.7 are similar to tables 4.4 and 4.5 and report the results for the maximum producible output and the inverse of technical efficiency scores respectively. The correlation between the bias-corrected estimates of the maximum producible output is now only 0.21. The bias corrected estimates originated from the extended DEA model are all

[1] It is possible the observed output to be higher than the upper bound of the 95% confidence interval.

higher than the bias-corrected estimates generated through the bootstrap conditional on the socioeconomic characteristics.

In Table 4.8 we compare the alternative procedures for 6 selected school districts. Three of them- Bridgeport, Hartford and New Haven- are large urban districts and the other three –Avon, Glastonbury, and Simsbury- are, by contrast, affluent suburban districts. The alternative procedures yield comparable results for the suburban districts. On the other hand, large differences can be found for the poorer urban districts. The conventional bootstrap approach suggests that these urban districts can achieve higher SAT scores. This reinforces the belief for example, that Hartford is a poorly managed district. Our results suggest however that given the socioeconomic disadvantages in the homelife of an average student, this district is performing even better than what could be expected. The optimum SAT score from the bias corrected column is lower than the observed one.

4.5 Conclusion

Researchers use bootstrap as a way to generate empirical distributions of efficiency scores from DEA models. The existing bootstrap procedures assume that all DMUs have the same probability of getting an efficiency score. However, there are usually systematic factors that contribute to efficiency. In this chapter an alternative bootstrap technique is developed that generates distributions of efficiency scores conditional on such factors. One might argue in favor of including them in the DEA model, but this might violate the free disposability assumptions. Also, as we have shown with the empirical application, the two-level method leads to results that are more consistent with these systematic factors than the conventional method.

Table 4.1: Summary Statistics of the Data

Description	Mean	St. Deviation
Output:		
Average SAT Score	1017.59	70.45
Inputs:		
(i) regular instruction teachers per pupil	0.06213	0.00697
(ii) special education teachers per pupil	0.00944	0.00195
(iii) support staff (like counselors) per pupil	0.00557	0.00147
(iv) administrators per pupil	0.00531	0.00168
(v) computers per pupil	0.13219	0.04563
(vi) total instruction hours during the school year	972.03	29.13
Socio-economic variables:		
(i) proportion of minorities in the school district population	14.06 %	20.01
(ii) proportion of adults with a college degree in the school district population	41.03 %	19.73

Sourse: Profiles of our Schools. Condition of Education in Connecticut 1995-1996, Connecticut State Board of Education, February 1997.

Table 4.2: Output Oriented Technical Efficiency

		DEA without socio-economic factors		DEA with socio-economic factors	
School District	**SAT**	ϕ^0	SAT* = SAT * ϕ^0	ϕ^1	SAT* = SAT * ϕ^1
ANSONIA	962	1.0326	993.40	1.0326	993.40
AVON	1133	1.0000	1133.00	1.0000	1133.00
BERLIN	987	1.1321	1117.40	1.0732	1059.22
BETHEL	1051	1.0644	1118.65	1.0513	1104.89
BLOOMFIELD	886	1.2306	1090.32	1.1583	1026.22
BOLTON	1078	1.0000	1078.00	1.0000	1078.00
BRANFORD	1027	1.0000	1027.00	1.0000	1027.00
BRIDGEPORT	782	1.3924	1088.83	1.0000	782.00
BRISTOL	1004	1.0620	1066.21	1.0000	1004.00
BROOKFIELD	1053	1.0343	1089.08	1.0343	1089.08
CANTON	1061	1.0187	1080.79	1.0000	1061.00
CHESHIRE	1063	1.0638	1130.77	1.0570	1123.58
CLINTON	1028	1.0787	1108.93	1.0606	1090.31
COLCHESTER	1019	1.0000	1019.00	1.0000	1019.00
COVENTRY	1031	1.0831	1116.64	1.0263	1058.08
CROMWELL	987	1.1236	1108.99	1.0997	1085.36
DANBURY	956	1.1018	1053.31	1.0795	1032.00

DARIEN	1121	1.0000	1121.00	1.0000	1121.00
DERBY	914	1.0031	916.80	1.0000	914.00
EAST GRANBY	1065	1.0000	1065.00	1.0000	1065.00
EAST HADDAM	1036	1.0386	1076.00	1.0294	1066.46
EAST HAMPTON	997	1.0537	1050.54	1.0393	1036.21
EAST HARTFORD	973	1.0316	1003.78	1.0009	973.87
EAST HAVEN	908	1.1324	1028.25	1.0262	931.76
EAST LYME	1068	1.0346	1104.95	1.0116	1080.44
EAST WINDSOR	963	1.1192	1077.79	1.0000	963.00
ELLINGTON	1050	1.0731	1126.72	1.0418	1093.88
ENFIELD	1011	1.0232	1034.45	1.0000	1011.00
FAIRFIELD	1068	1.0564	1128.25	1.0448	1115.82
FARMINGTON	1092	1.0000	1092.00	1.0000	1092.00
GLASTONBURY	1074	1.0161	1091.25	1.0000	1074.00
GRANBY	1097	1.0150	1113.50	1.0033	1100.64
GREENWICH	1078	1.0588	1141.41	1.0465	1128.13
GRISWOLD	964	1.1114	1071.36	1.0000	964.00
GROTON	1037	1.0341	1072.41	1.0005	1037.47
GUILFORD	1073	1.0354	1111.01	1.0326	1107.95
HAMDEN	962	1.1372	1093.96	1.1339	1090.85
HARTFORD	777	1.3251	1029.62	1.0000	777.00
KILLINGLY	992	1.1189	1109.93	1.0000	992.00

LEBANON	1036	1.1016	1141.30	1.0000	1036.00
LEDYARD	1045	1.0869	1135.84	1.0466	1093.73
LITCHFIELD	1077	1.0330	1112.59	1.0000	1077.00
MADISON	1087	1.0356	1125.66	1.0000	1087.00
MANCHESTER	1000	1.0740	1073.99	1.0570	1056.96
MERIDEN	946	1.0940	1034.96	1.0000	946.00
MIDDLETOWN	961	1.1481	1103.35	1.0823	1040.10
MILFORD	973	1.1340	1103.43	1.0971	1067.45
MONROE	1007	1.0317	1038.93	1.0293	1036.53
MONTVILLE	1007	1.0900	1097.64	1.0309	1038.12
NAUGATUCK	928	1.0134	940.42	1.0000	928.00
NEW BRITAIN	909	1.1070	1006.26	1.0771	979.11
NEW CANAAN	1129	1.0029	1132.32	1.0000	1129.00
NEW FAIRFIELD	1048	1.0000	1048.00	1.0000	1048.00
NEW HAVEN	789	1.3568	1070.52	1.2461	983.17
NEW INGTON	1032	1.0769	1111.35	1.0231	1055.85
NEW LONDON	893	1.2541	1119.93	1.1346	1013.22
NEW MILFORD	1058	1.0000	1058.00	1.0000	1058.00
NEWTOWN	1065	1.0075	1072.94	1.0000	1065.00
NORTH BRANFORD	1002	1.0583	1060.37	1.0000	1002.00
NORTH HAVEN	999	1.1209	1119.81	1.0837	1082.57
NORTH STONINGTON	1061	1.0281	1090.83	1.0000	1061.00

NORWALK	943	1.1452	1079.96	1.0897	1027.56
OLD SAYBROOK	1052	1.0304	1084.03	1.0296	1083.18
PLAINFIELD	975	1.1202	1092.16	1.0000	975.00
PLAINVILLE	990	1.0630	1052.33	1.0000	990.00
PLYMOUTH	983	1.0871	1068.64	1.0122	994.95
PORTLAND	1033	1.1017	1138.05	1.0605	1095.46
PUTNAM	986	1.1008	1085.43	1.0320	1017.60
RIDGEFIELD	1130	1.0000	1130.00	1.0000	1130.00
ROCKY HILL	1011	1.0819	1093.81	1.0676	1079.37
SEYMOUR	975	1.0000	975.00	1.0000	975.00
SHELTON	982	1.0004	982.38	1.0000	982.00
SIMSBURY	1140	1.0000	1140.00	1.0000	1140.00
SOMERS	1017	1.0588	1076.85	1.0000	1017.00
SOUTHINGTON	993	1.1260	1118.12	1.0998	1092.09
SOUTH WINDSOR	1051	1.0210	1073.04	1.0210	1073.04
STAFFORD	1080	1.0000	1080.00	1.0000	1080.00
STAMFORD	950	1.1514	1093.83	1.0887	1034.30
STONINGTON	1034	1.0108	1045.20	1.0000	1034.00
STRATFORD	948	1.1535	1093.52	1.1127	1054.85
SUFFIELD	1026	1.0369	1063.90	1.0339	1060.83
THOMASTON	937	1.0000	937.00	1.0000	937.00
THOMPSON	1060	1.0000	1060.00	1.0000	1060.00

TOLLAND	1089	1.0000	1089.00	1.0000	1089.00
TORRINGTON	970	1.0712	1039.08	1.0519	1020.31
TRUMBULL	1029	1.0327	1062.70	1.0327	1062.70
VERNON	1036	1.0712	1109.76	1.0499	1087.73
WALLINGFORD	968	1.0077	975.43	1.0000	968.00
WATERBURY	856	1.0000	856.00	1.0000	856.00
WATERFORD	1033	1.1048	1141.30	1.0603	1095.31
WATERTOWN	1000	1.0227	1022.69	1.0159	1015.87
WESTBROOK	1024	1.0489	1074.07	1.0489	1074.05
WEST HARTFORD	1070	1.0000	1070.00	1.0000	1070.00
WEST HAVEN	935	1.0000	935.00	1.0000	935.00
WESTON	1110	1.0011	1111.25	1.0000	1110.00
WESTPORT	1139	1.0023	1141.63	1.0000	1139.00
WETHERSFIELD	1012	1.1099	1123.21	1.0990	1112.15
WILTON	1142	1.0000	1142.00	1.0000	1142.00
WINDHAM	992	1.1422	1133.03	1.0557	1047.23
WINDSOR	989	1.1086	1096.42	1.0906	1078.57
WINDSOR LOCKS	1008	1.0000	1008.00	1.0000	1008.00
WOLCOTT	942	1.1147	1050.06	1.0475	986.79
RD1	1005	1.0883	1093.69	1.0589	1064.17
RD4	1053	1.0814	1138.68	1.0518	1107.55
RD5	1080	1.0471	1130.87	1.0375	1120.53

RD6	1116	1.0201	1138.41	1.0000	1116.00
RD7	1089	1.0384	1130.79	1.0000	1089.00
RD8	1096	1.0385	1138.24	1.0006	1096.67
RD9	1090	1.0222	1114.22	1.0000	1090.00
RD10	1049	1.0827	1135.79	1.0460	1097.25
RD11	1014	1.1262	1142.00	1.0797	1094.79
RD12	1010	1.0997	1110.66	1.0871	1097.99
RD13	1039	1.0588	1100.05	1.0515	1092.47
RD14	1002	1.0000	1002.00	1.0000	1002.00
RD15	1028	1.0572	1086.83	1.0516	1081.01
RD17	1049	1.0702	1122.67	1.0514	1102.94
RD18	1117	1.0108	1129.03	1.0000	1117.00
RD19	1119	1.0000	1119.00	1.0000	1119.00
Average	1017.59	1.0649	1080.13	1.0314	1048.35

Table 4.3: OLS Results

	Parameter Estimate	Standard Error	t-ratio
Intercept	1.0676	0.0133	80.1390
BA	-0.0008	0.0003	-3.0810
Minority	0.0021	0.0003	8.3550

$R^2 = 52.12\%$

n = 118

Table 4.4: Bootstrap results for Maximum Producible Output

School District	y_i^f	$\bar{\hat{y}}_i^f$	$se_B(y_i^f)$	$bias_B(y_i^f)$	$\hat{\hat{y}}_i^f$	95% lower limit	95% upper limit
ANSONIA	993.40	1046.79	94.21	53.39	940.01	858.90	956.74
AVON	1133.00	1120.33	119.05	-12.67	1145.67	951.76	1053.80
BERLIN	1117.40	1351.44	82.88	234.04	883.36	790.91	871.26
BETHEL	1118.65	1231.22	101.40	112.57	1006.08	945.13	1035.01
BLOOMFIELD	1090.32	1455.13	59.61	364.81	725.51	672.40	740.81
BOLTON	1078.00	1065.53	110.82	-12.47	1090.47	963.03	1077.24
BRANFORD	1027.00	1006.73	103.07	-20.27	1047.27	928.50	1008.51
BRIDGEPORT	1088.83	1610.31	42.78	521.49	567.34	529.53	569.55
BRISTOL	1066.21	1170.80	93.79	104.59	961.63	861.42	953.44
BROOKFIELD	1089.08	1146.81	95.97	57.72	1031.36	871.56	965.03
CANTON	1080.79	1106.62	110.59	25.83	1054.96	965.87	1082.90
CHESHIRE	1130.77	1254.35	96.93	123.58	1007.20	911.63	1014.52
CLINTON	1108.93	1250.34	93.21	141.41	967.52	842.98	924.62
COLCHESTER	1019.00	999.14	109.08	-19.86	1038.86	830.15	915.98
COVENTRY	1116.64	1269.82	94.24	153.18	963.46	908.12	1002.71
CROMWELL	1108.99	1326.28	82.50	217.29	891.70	778.37	859.62
DANBURY	1053.31	1231.93	78.66	178.63	874.68	786.45	858.72

DARIEN	1121.00	1105.84	118.11	-15.16	1136.16	1077.56	1198.61
DERBY	916.80	907.62	95.85	-9.18	925.98	825.29	920.97
EAST GRANBY	1065.00	1052.69	109.90	-12.31	1077.31	911.98	1009.20
EAST HADDAM	1076.00	1145.76	96.46	69.77	1006.23	906.45	996.10
EAST HAMPTON	1050.54	1146.77	95.78	96.23	954.32	880.85	983.02
EAST HARTFORD	1003.78	1054.28	98.26	50.51	953.27	838.58	939.94
EAST HAVEN	1028.25	1244.84	71.30	216.59	811.66	692.24	761.35
EAST LYME	1104.95	1163.49	107.43	58.54	1046.41	871.13	972.21
EAST WINDSOR	1077.79	1285.17	73.97	207.38	870.41	848.63	916.26
ELLINGTON	1126.72	1259.87	95.48	133.15	993.57	881.63	967.88
ENFIELD	1034.45	1065.86	95.97	31.41	1003.03	1063.59	1155.93
FAIRFIELD	1128.25	1230.30	97.03	102.05	1026.20	960.07	1055.47
FARMINGTON	1092.00	1077.17	115.10	-14.83	1106.83	1010.31	1111.77
GLASTONBURY	1091.25	1106.25	105.52	15.01	1076.24	1058.10	1164.21
GRANBY	1113.50	1131.73	115.83	18.23	1095.26	1013.42	1127.12
GREENWICH	1141.41	1259.84	104.05	118.44	1022.97	919.66	1028.95
GRISWOLD	1071.36	1270.97	78.25	199.62	871.74	775.24	848.79
GROTON	1072.41	1131.40	103.81	58.99	1013.42	925.63	1025.09
GUILFORD	1111.01	1170.33	101.76	59.32	1051.69	873.73	960.02
HAMDEN	1093.96	1332.57	81.50	238.62	855.34	764.43	848.08
HARTFORD	1029.62	1467.84	45.67	438.22	591.39	493.82	531.30
KILLINGLY	1109.93	1326.33	78.62	216.40	893.53	814.22	894.20

LEBANON	1141.30	1331.46	90.77	190.16	951.14	831.83	905.73
LEDYARD	1135.84	1298.07	91.04	162.23	973.61	915.89	1020.61
LITCHFIELD	1112.59	1170.53	108.79	57.94	1054.65	972.00	1086.31
MADISON	1125.66	1189.14	111.71	63.47	1062.19	789.62	881.38
MANCHESTER	1073.99	1204.68	89.55	130.70	943.29	810.00	894.62
MERIDEN	1034.96	1192.09	79.73	157.13	877.82	841.88	914.88
MIDDLETOWN	1103.35	1364.11	69.40	260.76	842.60	743.89	807.23
MILFORD	1103.43	1336.34	83.86	232.92	870.51	749.16	830.36
MONROE	1038.93	1085.23	99.80	46.30	992.63	881.82	976.80
MONTVILLE	1097.64	1260.31	89.50	162.66	934.98	809.59	901.80
NAUGATUCK	940.42	953.95	96.24	13.53	926.89	688.89	776.00
NEW BRITAIN	1006.26	1180.85	74.43	174.59	831.66	725.57	794.68
NEW CANAAN	1132.32	1130.31	110.81	-2.00	1134.32	992.57	1114.39
NEW FAIRFIELD	1048.00	1038.63	109.64	-9.37	1057.37	882.71	982.71
NEW HAVEN	1070.52	1551.98	44.88	481.46	589.06	585.26	628.78
NEW INGTON	1111.35	1253.83	94.27	142.48	968.87	782.16	878.40
NEW LONDON	1119.93	1521.99	60.16	402.06	717.87	649.55	707.22
NEW MILFORD	1058.00	1045.88	113.49	-12.12	1070.12	973.04	1085.48
NEWTOWN	1072.94	1062.07	106.08	-10.87	1083.80	947.94	1034.43
NORTH BRANFORD	1060.37	1163.10	91.10	102.73	957.63	943.68	1037.43
NORTH HAVEN	1119.81	1337.09	82.08	217.28	902.53	813.70	895.21
NORTH STONINGTON	1090.83	1135.85	101.48	45.02	1045.81	881.12	961.06

NORWALK	1079.96	1324.23	75.01	244.27	835.69	794.39	872.09
OLD SAYBROOK	1084.03	1136.49	106.56	52.46	1031.57	981.19	1093.00
PLAINFIELD	1092.16	1297.47	84.48	205.31	886.86	820.68	884.33
PLAINVILLE	1052.33	1168.36	88.07	116.03	936.30	914.31	1004.74
PLYMOUTH	1068.64	1217.84	88.29	149.20	919.43	804.79	875.93
PORTLAND	1138.05	1327.13	88.96	189.09	948.96	799.44	879.54
PUTNAM	1085.43	1262.82	86.27	177.38	908.05	855.74	928.26
RIDGEFIELD	1130.00	1117.20	118.10	-12.80	1142.80	1044.82	1167.03
ROCKY HILL	1093.81	1249.39	92.21	155.58	938.22	822.31	914.44
SEYMOUR	975.00	961.07	104.62	-13.93	988.93	917.57	1025.80
SHELTON	982.38	971.45	98.39	-10.93	993.31	903.34	1001.99
SIMSBURY	1140.00	1121.72	119.27	-18.28	1158.28	1122.56	1226.06
SOMERS	1076.85	1184.51	93.56	107.67	969.18	855.24	951.39
SOUTHINGTON	1118.12	1345.99	76.10	227.88	890.24	839.59	908.63
SOUTH WINDSOR	1073.04	1107.23	105.66	34.19	1038.85	941.22	1050.51
STAFFORD	1080.00	1063.06	109.94	-16.94	1096.94	950.02	1064.94
STAMFORD	1093.83	1352.51	74.73	258.68	835.15	770.08	837.30
STONINGTON	1045.20	1057.85	101.36	12.65	1032.55	859.26	950.40
STRATFORD	1093.52	1354.55	73.48	261.02	832.50	760.49	827.79
SUFFIELD	1063.90	1129.27	94.78	65.37	998.53	927.46	997.22
THOMASTON	937.00	926.04	100.81	-10.96	947.96	808.64	902.57
THOMPSON	1060.00	1041.42	108.03	-18.58	1078.58	948.20	1062.51

TOLLAND	1089.00	1071.69	116.83	-17.31	1106.31	935.61	1055.22
TORRINGTON	1039.08	1159.62	87.01	120.54	918.54	823.64	900.52
TRUMBULL	1062.70	1113.14	101.25	50.44	1012.26	877.52	967.56
VERNON	1109.76	1246.84	91.83	137.08	972.68	887.38	979.18
WALLINGFORD	975.43	964.82	98.52	-10.61	986.04	917.23	1011.08
WATERBURY	856.00	841.60	85.76	-14.40	870.40	771.92	844.24
WATERFORD	1141.30	1336.86	85.60	195.57	945.73	864.94	937.24
WATERTOWN	1022.69	1049.65	102.20	26.97	995.72	885.83	988.60
WESTBROOK	1074.07	1154.49	102.33	80.42	993.64	851.63	954.50
WEST HARTFORD	1070.00	1063.20	112.99	-6.80	1076.80	955.51	1063.13
WEST HAVEN	935.00	922.58	97.75	-12.42	947.42	838.74	943.56
WESTON	1111.25	1103.39	114.28	-7.85	1119.10	930.65	1040.38
WESTPORT	1141.63	1131.92	114.74	-9.71	1151.34	1108.42	1226.37
WETHERSFIELD	1123.21	1319.60	86.76	196.39	926.82	822.84	900.73
WILTON	1142.00	1127.25	114.12	-14.75	1156.75	1010.58	1106.26
WINDHAM	1133.03	1386.68	82.28	253.65	879.37	781.69	868.09
WINDSOR	1096.42	1292.39	81.94	195.97	900.45	806.43	887.37
WINDSOR LOCKS	1008.00	995.89	97.90	-12.11	1020.11	906.59	993.86
WOLCOTT	1050.06	1241.13	74.80	191.06	859.00	755.95	814.78
RD1	1093.69	1254.32	94.79	160.63	933.07	840.14	940.33
RD4	1138.68	1294.57	91.48	155.90	982.78	933.77	1027.92
RD5	1130.87	1213.29	105.84	82.42	1048.45	997.03	1103.15

RD6	1138.41	1164.30	111.30	25.89	1112.52	1052.94	1167.29
RD7	1130.79	1192.43	104.38	61.65	1069.14	1026.00	1117.54
RD8	1138.24	1206.90	106.12	68.67	1069.57	972.19	1071.00
RD9	1114.22	1144.58	108.67	30.36	1083.86	920.52	1019.48
RD10	1135.79	1287.03	90.48	151.24	984.55	925.34	1006.40
RD11	1142.00	1372.68	82.65	230.68	911.32	834.87	916.05
RD12	1110.66	1293.14	86.36	182.47	928.19	860.20	944.70
RD13	1100.05	1206.65	95.47	106.60	993.45	920.29	1003.56
RD14	1002.00	987.63	100.64	-14.37	1016.37	895.95	983.76
RD15	1086.83	1183.86	99.91	97.03	989.81	834.53	932.69
RD17	1122.67	1254.76	93.91	132.09	990.59	908.33	985.92
RD18	1129.03	1133.44	109.74	4.41	1124.62	1078.49	1187.20
RD19	1119.00	1097.73	126.44	-21.27	1140.27	1024.12	1138.05
Average	1080.13	1185.73	94.91	-105.59	974.54		

y_i^f = maximum producible output from the DEA model

$\bar{y}_i^f$ = bootstrap average of the maximum producible output

$se_B(y_i^f)$ = standard deviation of the maximum producible output

$bias_B(y_i^f)$ = bias of the maximum producible output

$\hat{y}_i^f$ = bias-corrected estimate of the maximum producible output

Table 4.5: Bootstrap results for ϕ scores

School District	ϕ_i	$\bar{\hat{\phi}}_i$	$se_B(\hat{\phi}_i)$	$bias_B(\hat{\phi}_i)$	$\hat{\hat{\phi}}_i$	95% lower limit	95% upper limit
ANSONIA	1.0326	1.0334	0.0518	0.0008	1.0319	0.9364	1.0299
AVON	1.0000	1.0015	0.0521	0.0015	0.9985	0.9093	0.9977
BERLIN	1.1321	1.1358	0.0530	0.0037	1.1284	1.0270	1.1329
BETHEL	1.0644	1.0703	0.0534	0.0059	1.0585	0.9735	1.0659
BLOOMFIELD	1.2306	1.2322	0.0513	0.0016	1.2290	1.1451	1.2261
BOLTON	1.0000	1.0019	0.0507	0.0019	0.9981	0.9033	0.9977
BRANFORD	1.0000	1.0062	0.0479	0.0062	0.9938	0.9190	1.0049
BRIDGEPORT	1.3924	1.3964	0.0523	0.0041	1.3883	1.2984	1.3951
BRISTOL	1.0620	1.0671	0.0527	0.0052	1.0568	0.9783	1.0619
BROOKFIELD	1.0343	1.0375	0.0488	0.0033	1.0310	0.9535	1.0334
CANTON	1.0187	1.0210	0.0536	0.0023	1.0163	0.9278	1.0183
CHESHIRE	1.0638	1.0642	0.0523	0.0004	1.0633	0.9718	1.0593
CLINTON	1.0787	1.0832	0.0516	0.0044	1.0743	0.9791	1.0785
COLCHESTER	1.0000	1.0056	0.0517	0.0056	0.9944	0.9082	1.0013
COVENTRY	1.0831	1.0857	0.0531	0.0026	1.0804	0.9857	1.0834
CROMWELL	1.1236	1.1284	0.0522	0.0048	1.1188	1.0242	1.1258
DANBURY	1.1018	1.1028	0.0500	0.0011	1.1007	1.0155	1.0999

DARIEN	1.0000	1.0027	0.0521	0.0027	0.9973	0.9076	1.0002
DERBY	1.0031	1.0071	0.0521	0.0040	0.9990	0.9088	1.0042
EAST GRANBY	1.0000	1.0019	0.0509	0.0019	0.9981	0.9086	0.9980
EAST HADDAM	1.0386	1.0395	0.0504	0.0009	1.0377	0.9541	1.0362
EAST HAMPTON	1.0537	1.0542	0.0535	0.0005	1.0532	0.9655	1.0506
EAST HARTFORD	1.0316	1.0329	0.0546	0.0013	1.0303	0.9404	1.0294
EAST HAVEN	1.1324	1.1361	0.0501	0.0036	1.1288	1.0479	1.1331
EAST LYME	1.0346	1.0375	0.0536	0.0029	1.0317	0.9428	1.0319
EAST WINDSOR	1.1192	1.1226	0.0472	0.0034	1.1158	1.0357	1.1202
ELLINGTON	1.0731	1.0776	0.0507	0.0045	1.0685	0.9702	1.0766
ENFIELD	1.0232	1.0274	0.0483	0.0042	1.0190	0.9388	1.0243
FAIRFIELD	1.0564	1.0607	0.0503	0.0043	1.0521	0.9670	1.0562
FARMINGTON	1.0000	1.0028	0.0513	0.0028	0.9972	0.9054	0.9994
GLASTONBURY	1.0161	1.0216	0.0494	0.0055	1.0105	0.9299	1.0198
GRANBY	1.0150	1.0176	0.0535	0.0025	1.0125	0.9142	1.0133
GREENWICH	1.0588	1.0572	0.0542	-0.0016	1.0604	0.9573	1.0541
GRISWOLD	1.1114	1.1108	0.0505	-0.0006	1.1120	1.0249	1.1090
GROTON	1.0341	1.0356	0.0525	0.0015	1.0327	0.9342	1.0358
GUILFORD	1.0354	1.0394	0.0504	0.0040	1.0315	0.9455	1.0364
HAMDEN	1.1372	1.1395	0.0547	0.0024	1.1348	1.0312	1.1348
HARTFORD	1.3251	1.3278	0.0515	0.0027	1.3224	1.2359	1.3253
KILLINGLY	1.1189	1.1200	0.0495	0.0011	1.1178	1.0333	1.1162

LEBANON	1.1016	1.1042	0.0523	0.0025	1.0991	1.0064	1.1038
LEDYARD	1.0869	1.0901	0.0518	0.0032	1.0837	1.0000	1.0859
LITCHFIELD	1.0330	1.0349	0.0539	0.0018	1.0312	0.9359	1.0310
MADISON	1.0356	1.0374	0.0546	0.0019	1.0337	0.9325	1.0314
MANCHESTER	1.0740	1.0773	0.0512	0.0033	1.0706	0.9858	1.0742
MERIDEN	1.0940	1.0989	0.0492	0.0049	1.0891	1.0117	1.0989
MIDDLETOWN	1.1481	1.1490	0.0475	0.0009	1.1472	1.0636	1.1460
MILFORD	1.1340	1.1383	0.0548	0.0043	1.1298	1.0312	1.1343
MONROE	1.0317	1.0365	0.0516	0.0048	1.0269	0.9357	1.0328
MONTVILLE	1.0900	1.0928	0.0528	0.0027	1.0873	0.9980	1.0894
NAUGATUCK	1.0134	1.0154	0.0529	0.0020	1.0114	0.9218	1.0106
NEW BRITAIN	1.1070	1.1107	0.0495	0.0037	1.1033	1.0264	1.1067
NEW CANAAN	1.0029	1.0031	0.0500	0.0002	1.0028	0.9198	0.9992
NEW FAIRFIELD	1.0000	1.0004	0.0524	0.0004	0.9996	0.9048	0.9980
NEW HAVEN	1.3568	1.3626	0.0518	0.0058	1.3510	1.2633	1.3593
NEW INGTON	1.0769	1.0789	0.0533	0.0020	1.0749	0.9889	1.0717
NEW LONDON	1.2541	1.2559	0.0528	0.0018	1.2523	1.1661	1.2541
NEW MILFORD	1.0000	1.0015	0.0536	0.0015	0.9985	0.9132	0.9952
NEWTOWN	1.0075	1.0163	0.0491	0.0089	0.9986	0.9282	1.0122
NORTH BRANFORD	1.0583	1.0606	0.0508	0.0024	1.0559	0.9728	1.0555
NORTH HAVEN	1.1209	1.1245	0.0513	0.0036	1.1174	1.0290	1.1234
NORTH STONINGTON	1.0281	1.0314	0.0494	0.0033	1.0249	0.9409	1.0298

NORWALK	1.1452	1.1502	0.0522	0.0049	1.1403	1.0665	1.1452
OLD SAYBROOK	1.0304	1.0317	0.0538	0.0012	1.0292	0.9299	1.0271
PLAINFIELD	1.1202	1.1269	0.0521	0.0068	1.1134	1.0237	1.1245
PLAINVILLE	1.0630	1.0623	0.0502	-0.0007	1.0637	0.9776	1.0584
PLYMOUTH	1.0871	1.0924	0.0519	0.0053	1.0818	0.9872	1.0877
PORTLAND	1.1017	1.1046	0.0522	0.0029	1.0988	1.0060	1.1000
PUTNAM	1.1008	1.1047	0.0522	0.0038	1.0970	1.0096	1.1029
RIDGEFIELD	1.0000	1.0016	0.0518	0.0016	0.9984	0.9058	0.9988
ROCKY HILL	1.0819	1.0801	0.0537	-0.0018	1.0837	0.9906	1.0746
SEYMOUR	1.0000	1.0029	0.0533	0.0029	0.9971	0.9091	0.9996
SHELTON	1.0004	1.0026	0.0499	0.0022	0.9982	0.9131	0.9997
SIMSBURY	1.0000	1.0041	0.0505	0.0041	0.9959	0.8978	1.0013
SOMERS	1.0588	1.0599	0.0521	0.0010	1.0578	0.9726	1.0551
SOUTHINGTON	1.1260	1.1282	0.0487	0.0022	1.1238	1.0480	1.1231
SOUTH WINDSOR	1.0210	1.0217	0.0523	0.0007	1.0203	0.9243	1.0166
STAFFORD	1.0000	1.0040	0.0508	0.0040	0.9960	0.9192	0.9999
STAMFORD	1.1514	1.1550	0.0507	0.0036	1.1478	1.0656	1.1518
STONINGTON	1.0108	1.0119	0.0499	0.0011	1.0098	0.9285	1.0098
STRATFORD	1.1535	1.1575	0.0514	0.0040	1.1495	1.0665	1.1506
SUFFIELD	1.0369	1.0383	0.0482	0.0013	1.0356	0.9485	1.0343
THOMASTON	1.0000	1.0016	0.0535	0.0016	0.9984	0.9061	0.9992
THOMPSON	1.0000	1.0049	0.0509	0.0049	0.9951	0.9129	0.9997

TOLLAND	1.0000	1.0036	0.0537	0.0036	0.9964	0.9152	0.9983
TORRINGTON	1.0712	1.0753	0.0501	0.0041	1.0671	0.9759	1.0746
TRUMBULL	1.0327	1.0371	0.0515	0.0043	1.0284	0.9464	1.0334
VERNON	1.0712	1.0707	0.0509	-0.0005	1.0717	0.9788	1.0675
WALLINGFORD	1.0077	1.0171	0.0492	0.0094	0.9983	0.9255	1.0157
WATERBURY	1.0000	1.0047	0.0492	0.0047	0.9953	0.9145	1.0046
WATERFORD	1.1048	1.1078	0.0501	0.0029	1.1019	1.0173	1.1055
WATERTOWN	1.0227	1.0279	0.0529	0.0052	1.0175	0.9348	1.0232
WESTBROOK	1.0489	1.0539	0.0544	0.0050	1.0438	0.9572	1.0506
WEST HARTFORD	1.0000	1.0009	0.0527	0.0009	0.9991	0.9094	0.9932
WEST HAVEN	1.0000	1.0025	0.0525	0.0025	0.9975	0.9087	0.9968
WESTON	1.0011	1.0018	0.0523	0.0007	1.0005	0.9144	0.9954
WESTPORT	1.0023	1.0052	0.0498	0.0028	0.9995	0.9156	1.0023
WETHERSFIELD	1.1099	1.1154	0.0511	0.0055	1.1044	1.0092	1.1139
WILTON	1.0000	1.0028	0.0493	0.0028	0.9972	0.9139	1.0008
WINDHAM	1.1422	1.1455	0.0541	0.0033	1.1389	1.0503	1.1406
WINDSOR	1.1086	1.1103	0.0514	0.0017	1.1070	1.0233	1.1072
WINDSOR LOCKS	1.0000	1.0025	0.0481	0.0025	0.9975	0.9137	0.9981
WOLCOTT	1.1147	1.1207	0.0481	0.0060	1.1088	1.0333	1.1196
RD1	1.0883	1.0896	0.0565	0.0013	1.0869	0.9956	1.0827
RD4	1.0814	1.0828	0.0514	0.0014	1.0800	0.9963	1.0779
RD5	1.0471	1.0517	0.0523	0.0046	1.0425	0.9433	1.0510

RD6	1.0201	1.0248	0.0513	0.0047	1.0154	0.9309	1.0233
RD7	1.0384	1.0447	0.0500	0.0063	1.0321	0.9485	1.0421
RD8	1.0385	1.0417	0.0513	0.0032	1.0354	0.9436	1.0402
RD9	1.0222	1.0267	0.0509	0.0045	1.0177	0.9340	1.0251
RD10	1.0827	1.0881	0.0490	0.0053	1.0774	0.9984	1.0870
RD11	1.1262	1.1296	0.0510	0.0034	1.1229	1.0375	1.1258
RD12	1.0997	1.1019	0.0513	0.0022	1.0974	1.0001	1.0985
RD13	1.0588	1.0616	0.0509	0.0028	1.0559	0.9677	1.0580
RD14	1.0000	1.0035	0.0494	0.0035	0.9965	0.9146	0.9996
RD15	1.0572	1.0625	0.0540	0.0053	1.0519	0.9609	1.0583
RD17	1.0702	1.0723	0.0504	0.0021	1.0681	0.9814	1.0705
RD18	1.0108	1.0159	0.0500	0.0052	1.0056	0.9239	1.0113
RD19	1.0000	1.0049	0.0549	0.0049	0.9951	0.8928	1.0003
Average	1.0649	1.0680	0.0515	0.0030	1.0619		

$\phi_i = \phi$ from the DEA model

$\bar{\hat{\phi}}_i$ = bootstrap average of ϕ

$se_B(\phi_i)$ = standard deviation of ϕ

$bias_B(\phi_i)$ = bias of ϕ

$\hat{\hat{\phi}}_i$ = bias-corrected estimate of ϕ

Table 4.6: Bootstrap results for Maximum Producible Output (extended DEA model)

School District	y_i^f	$\bar{\hat{y}}_i^f$	$se_B(y_i^f)$	$bias_B(y_i^f)$	$\hat{\hat{y}}_i^f$	95% lower limit	95% upper limit
ANSONIA	993.40	1046.79	60.25	53.39	933.14	1004.85	1088.47
AVON	1133.00	1120.33	80.85	-12.67	1072.76	1089.77	1201.07
BERLIN	1059.22	1351.44	84.49	234.04	1002.36	1041.90	1163.44
BETHEL	1104.89	1231.22	81.03	112.57	1039.70	1108.50	1213.13
BLOOMFIELD	1026.22	1455.13	76.53	364.81	968.67	904.74	994.77
BOLTON	1078.00	1065.53	77.01	-12.47	1016.56	1094.43	1194.90
BRANFORD	1027.00	1006.73	84.02	-20.27	964.86	930.12	1029.97
BRIDGEPORT	782.00	1610.31	55.85	521.49	739.09	746.99	815.15
BRISTOL	1004.00	1170.80	68.01	104.59	955.99	953.05	1036.75
BROOKFIELD	1089.08	1146.81	82.42	57.72	1024.12	1077.86	1179.90
CANTON	1061.00	1106.62	80.95	25.83	998.56	1045.53	1154.17
CHESHIRE	1123.58	1254.35	79.00	123.58	1060.92	1106.25	1199.96
CLINTON	1090.31	1250.34	79.74	141.41	1031.70	1075.59	1188.30
COLCHESTER	1019.00	999.14	74.51	-19.86	960.72	1023.01	1125.54
COVENTRY	1058.08	1269.82	79.15	153.18	1004.01	1009.12	1118.46
CROMWELL	1085.36	1326.28	82.94	217.29	1024.09	1046.31	1153.39
DANBURY	1032.00	1231.93	79.43	178.63	976.96	990.37	1103.02

DARIEN	1121.00	1105.84	88.25	-15.16	1059.57	1103.36	1225.54
DERBY	914.00	907.62	65.96	-9.18	865.25	781.78	865.56
EAST GRANBY	1065.00	1052.69	76.11	-12.31	1009.20	1071.02	1174.12
EAST HADDAM	1066.46	1145.76	75.08	69.77	1010.61	974.86	1062.42
EAST HAMPTON	1036.21	1146.77	73.30	96.23	977.58	1053.46	1147.36
EAST HARTFORD	973.87	1054.28	74.15	50.51	919.20	885.60	974.05
EAST HAVEN	931.76	1244.84	64.13	216.59	884.08	933.75	1018.02
EAST LYME	1080.44	1163.49	81.83	58.54	1023.56	1062.95	1170.15
EAST WINDSOR	963.00	1285.17	71.89	207.38	909.70	905.89	998.00
ELLINGTON	1093.88	1259.87	79.48	133.15	1035.67	1080.64	1188.46
ENFIELD	1011.00	1065.86	74.68	31.41	954.77	1022.71	1119.16
FAIRFIELD	1115.82	1230.30	86.10	102.05	1050.83	1058.25	1189.27
FARMINGTON	1092.00	1077.17	78.90	-14.83	1036.95	1073.64	1190.90
GLASTONBURY	1074.00	1106.25	77.96	15.01	1014.44	1037.66	1136.07
GRANBY	1100.64	1131.73	82.66	18.23	1037.83	1080.29	1191.65
GREENWICH	1128.13	1259.84	80.49	118.44	1065.16	1113.31	1219.52
GRISWOLD	964.00	1270.97	71.52	199.62	913.37	895.12	983.07
GROTON	1037.47	1131.40	78.92	58.99	977.54	1032.97	1135.09
GUILFORD	1107.95	1170.33	76.71	59.32	1044.97	1146.56	1238.17
HAMDEN	1090.85	1332.57	84.11	238.62	1024.45	1093.22	1201.08
HARTFORD	777.00	1467.84	59.16	438.22	732.48	720.78	793.78
KILLINGLY	992.00	1326.33	72.18	216.40	937.98	930.98	1025.52

LEBANON	1036.00	1331.46	79.04	190.16	977.62	1019.79	1130.78
LEDYARD	1093.73	1298.07	74.40	162.23	1039.05	881.07	959.76
LITCHFIELD	1077.00	1170.53	75.63	57.94	1018.01	1087.33	1183.87
MADISON	1087.00	1189.14	82.76	63.47	1024.39	1105.05	1211.58
MANCHESTER	1056.96	1204.68	79.48	130.70	999.99	982.27	1085.43
MERIDEN	946.00	1192.09	67.53	157.13	894.36	953.25	1043.54
MIDDLETOWN	1040.10	1364.11	76.33	260.76	981.29	990.04	1079.62
MILFORD	1067.45	1336.34	78.55	232.92	1005.55	1067.11	1170.11
MONROE	1036.53	1085.23	78.49	46.30	977.52	974.14	1069.48
MONTVILLE	1038.12	1260.31	70.70	162.66	987.46	1040.60	1127.27
NAUGATUCK	928.00	953.95	74.19	13.53	875.29	886.56	995.27
NEW BRITAIN	979.11	1180.85	70.34	174.59	924.79	961.81	1058.23
NEW CANAAN	1129.00	1130.31	79.28	-2.00	1066.84	1136.81	1246.22
NEW FAIRFIELD	1048.00	1038.63	79.48	-9.37	991.04	1026.35	1131.07
NEW HAVEN	983.17	1551.98	70.37	481.46	928.35	968.13	1059.85
NEW INGTON	1055.85	1253.83	78.65	142.48	999.69	1019.34	1125.18
NEW LONDON	1013.22	1521.99	75.31	402.06	960.08	990.80	1094.50
NEW MILFORD	1058.00	1045.88	77.05	-12.12	1001.03	1055.15	1158.55
NEWTOWN	1065.00	1062.07	78.78	-10.87	1011.90	1050.38	1156.71
NORTH BRANFORD	1002.00	1163.10	73.02	102.73	944.15	917.36	1007.47
NORTH HAVEN	1082.57	1337.09	80.21	217.28	1024.38	1020.46	1117.68
NORTH STONINGTON	1061.00	1135.85	80.28	45.02	999.23	1045.73	1139.59

NORWALK	1027.56	1324.23	77.63	244.27	967.12	1021.39	1132.69
OLD SAYBROOK	1083.18	1136.49	84.40	52.46	1019.14	982.07	1084.33
PLAINFIELD	975.00	1297.47	70.36	205.31	921.48	911.20	997.10
PLAINVILLE	990.00	1168.36	70.79	116.03	933.04	926.59	1011.88
PLYMOUTH	994.95	1217.84	70.95	149.20	940.72	980.62	1072.20
PORTLAND	1095.46	1327.13	85.35	189.09	1031.85	1075.21	1187.88
PUTNAM	1017.60	1262.82	75.12	177.38	965.20	950.31	1045.13
RIDGEFIELD	1130.00	1117.20	80.92	-12.80	1065.36	1134.71	1240.13
ROCKY HILL	1079.37	1249.39	75.61	155.58	1023.84	996.73	1073.77
SEYMOUR	975.00	961.07	73.93	-13.93	918.75	920.03	1012.08
SHELTON	982.00	971.45	72.92	-10.93	927.84	969.47	1071.06
SIMSBURY	1140.00	1121.72	83.39	-18.28	1080.56	1078.17	1182.35
SOMERS	1017.00	1184.51	77.27	107.67	962.70	936.06	1032.23
SOUTHINGTON	1092.09	1345.99	80.99	227.88	1033.43	1042.91	1158.27
SOUTH WINDSOR	1073.04	1107.23	84.32	34.19	1011.65	982.64	1083.21
STAFFORD	1080.00	1063.06	76.83	-16.94	1023.98	1089.38	1185.98
STAMFORD	1034.30	1352.51	72.53	258.68	981.57	1040.60	1134.43
STONINGTON	1034.00	1057.85	78.93	12.65	972.58	1038.18	1139.36
STRATFORD	1054.85	1354.55	75.09	261.02	996.39	972.63	1064.18
SUFFIELD	1060.83	1129.27	74.94	65.37	1004.62	981.24	1067.82
THOMASTON	937.00	926.04	67.70	-10.96	879.74	892.46	969.63
THOMPSON	1060.00	1041.42	80.39	-18.58	999.17	1025.58	1140.67

TOLLAND	1089.00	1071.69	78.36	-17.31	1028.59	1084.44	1195.78
TORRINGTON	1020.31	1159.62	73.55	120.54	967.17	874.50	966.56
TRUMBULL	1062.70	1113.14	77.67	50.44	1007.22	1011.20	1114.15
VERNON	1087.73	1246.84	79.50	137.08	1024.53	993.77	1089.37
WALLINGFORD	968.00	964.82	75.48	-10.61	911.54	940.60	1037.75
WATERBURY	856.00	841.60	68.42	-14.40	803.33	839.58	926.58
WATERFORD	1095.31	1336.86	84.36	195.57	1030.67	1024.46	1144.18
WATERTOWN	1015.87	1049.65	79.30	26.97	958.76	963.53	1072.04
WESTBROOK	1074.05	1154.49	76.81	80.42	1019.04	1085.62	1193.84
WEST HARTFORD	1070.00	1063.20	79.80	-6.80	1012.75	1059.56	1162.52
WEST HAVEN	935.00	922.58	74.05	-12.42	876.55	925.35	1019.60
WESTON	1110.00	1103.39	81.61	-7.85	1045.45	1107.57	1203.93
WESTPORT	1139.00	1131.92	85.01	-9.71	1080.62	1088.03	1204.29
WETHERSFIELD	1112.15	1319.60	84.98	196.39	1043.47	1116.43	1236.36
WILTON	1142.00	1127.25	77.37	-14.75	1081.36	1143.84	1248.71
WINDHAM	1047.23	1386.68	79.07	253.65	988.49	1057.95	1157.88
WINDSOR	1078.57	1292.39	82.89	195.97	1022.04	1033.04	1137.57
WINDSOR LOCKS	1008.00	995.89	69.36	-12.11	955.02	992.84	1091.06
WOLCOTT	986.79	1241.13	72.30	191.06	931.17	990.29	1098.15
RD1	1064.17	1254.32	77.93	160.63	1004.12	917.21	1007.45
RD4	1107.55	1294.57	79.49	155.90	1047.29	1063.40	1153.91
RD5	1120.53	1213.29	85.43	82.42	1059.18	1031.33	1144.34

RD6	1116.00	1164.30	83.51	25.89	1054.37	1070.04	1183.81
RD7	1089.00	1192.43	75.14	61.65	1032.03	1022.43	1113.42
RD8	1096.67	1206.90	76.58	68.67	1036.20	1087.02	1175.68
RD9	1090.00	1144.58	74.42	30.36	1030.70	1080.51	1167.44
RD10	1097.25	1287.03	78.11	151.24	1037.98	1090.20	1193.20
RD11	1094.79	1372.68	78.36	230.68	1035.27	1015.96	1110.89
RD12	1097.99	1293.14	81.21	182.47	1037.60	1001.23	1104.93
RD13	1092.47	1206.65	78.51	106.60	1032.97	1066.45	1163.09
RD14	1002.00	987.63	72.70	-14.37	949.97	1000.05	1099.02
RD15	1081.01	1183.86	76.80	97.03	1026.05	1015.36	1113.58
RD17	1102.94	1254.76	77.65	132.09	1045.95	1021.85	1117.89
RD18	1117.00	1133.44	86.15	4.41	1052.50	1007.92	1113.50
RD19	1119.00	1097.73	82.61	-21.27	1054.44	1019.25	1113.16
Average	1048.35	1185.73	77.07	57.87	990.48		

y_i^f = maximum producible output from the DEA model

$\bar{y}_i^f$ = bootstrap average of the maximum producible output

$se_B(y_i^f)$ = standard deviation of the maximum producible output

$bias_B(y_i^f)$ = bias of the maximum producible output

$\hat{y}_i^f$ = bias-corrected estimate of the maximum producible output

Table 4.7: Bootstrap results for ϕ scores (extended DEA model)

School District	ϕ_i	$\bar{\hat{\phi}}_i$	$se_B(\hat{\phi}_i)$	$bias_B(\hat{\phi}_i)$	$\hat{\hat{\phi}}_i$	95% lower limit	95% upper limit
ANSONIA	1.0326	1.0336	0.0369	0.0009	1.0317	0.9988	1.0151
AVON	1.0000	1.0301	0.0422	0.0301	0.9699	0.9171	0.9577
BERLIN	1.0732	1.0310	0.0479	-0.0422	1.1153	1.0599	1.0977
BETHEL	1.0513	1.0334	0.0424	-0.0179	1.0692	1.0142	1.0555
BLOOMFIELD	1.1583	1.0319	0.0444	-0.1264	1.2846	1.2322	1.2677
BOLTON	1.0000	1.0322	0.0425	0.0322	0.9678	0.9145	0.9551
BRANFORD	1.0000	1.0350	0.0501	0.0350	0.9650	0.9061	0.9470
BRIDGEPORT	1.0000	1.0310	0.0411	0.0310	0.9690	0.9159	0.9535
BRISTOL	1.0000	1.0269	0.0391	0.0269	0.9731	0.9243	0.9591
BROOKFIELD	1.0343	1.0339	0.0447	-0.0003	1.0346	0.9794	1.0204
CANTON	1.0000	1.0336	0.0459	0.0336	0.9664	0.9143	0.9504
CHESHIRE	1.0570	1.0314	0.0411	-0.0256	1.0826	1.0289	1.0684
CLINTON	1.0606	1.0305	0.0431	-0.0301	1.0907	1.0383	1.0765
COLCHESTER	1.0000	1.0324	0.0430	0.0324	0.9676	0.9137	0.9528
COVENTRY	1.0263	1.0293	0.0456	0.0030	1.0232	0.9733	1.0074
CROMWELL	1.0997	1.0322	0.0454	-0.0674	1.1671	1.1135	1.1495
DANBURY	1.0795	1.0306	0.0460	-0.0489	1.1284	1.0745	1.1109

DARIEN	1.0000	1.0315	0.0474	0.0315	0.9685	0.9095	0.9511
DERBY	1.0000	1.0302	0.0422	0.0302	0.9698	0.9175	0.9552
EAST GRANBY	1.0000	1.0296	0.0417	0.0296	0.9704	0.9210	0.9548
EAST HADDAM	1.0294	1.0296	0.0412	0.0002	1.0292	0.9775	1.0173
EAST HAMPTON	1.0393	1.0319	0.0414	-0.0074	1.0468	0.9924	1.0314
EAST HARTFORD	1.0009	1.0321	0.0457	0.0312	0.9697	0.9159	0.9549
EAST HAVEN	1.0262	1.0288	0.0402	0.0026	1.0235	0.9734	1.0099
EAST LYME	1.0116	1.0301	0.0450	0.0185	0.9932	0.9401	0.9778
EAST WINDSOR	1.0000	1.0315	0.0448	0.0315	0.9685	0.9166	0.9553
ELLINGTON	1.0418	1.0302	0.0431	-0.0116	1.0534	1.0022	1.0364
ENFIELD	1.0000	1.0316	0.0438	0.0316	0.9684	0.9156	0.9553
FAIRFIELD	1.0448	1.0334	0.0471	-0.0114	1.0562	1.0019	1.0407
FARMINGTON	1.0000	1.0286	0.0425	0.0286	0.9714	0.9235	0.9547
GLASTONBURY	1.0000	1.0314	0.0424	0.0314	0.9686	0.9124	0.9534
GRANBY	1.0033	1.0325	0.0449	0.0292	0.9741	0.9217	0.9593
GREENWICH	1.0465	1.0315	0.0414	-0.0150	1.0615	1.0085	1.0472
GRISWOLD	1.0000	1.0299	0.0442	0.0299	0.9701	0.9152	0.9536
GROTON	1.0005	1.0329	0.0452	0.0325	0.9680	0.9129	0.9520
GUILFORD	1.0326	1.0319	0.0406	-0.0006	1.0332	0.9794	1.0203
HAMDEN	1.1339	1.0347	0.0457	-0.0992	1.2332	1.1720	1.2187
HARTFORD	1.0000	1.0327	0.0453	0.0327	0.9673	0.9138	0.9506
KILLINGLY	1.0000	1.0309	0.0428	0.0309	0.9691	0.9150	0.9572

LEBANON	1.0000	1.0323	0.0469	0.0323	0.9677	0.9144	0.9525
LEDYARD	1.0466	1.0281	0.0394	-0.0185	1.0652	1.0158	1.0507
LITCHFIELD	1.0000	1.0308	0.0408	0.0308	0.9692	0.9175	0.9569
MADISON	1.0000	1.0329	0.0459	0.0329	0.9671	0.9121	0.9497
MANCHESTER	1.0570	1.0308	0.0455	-0.0261	1.0831	1.0287	1.0686
MERIDEN	1.0000	1.0308	0.0414	0.0308	0.9692	0.9160	0.9546
MIDDLETOWN	1.0823	1.0321	0.0437	-0.0502	1.1325	1.0790	1.1176
MILFORD	1.0971	1.0329	0.0433	-0.0642	1.1613	1.1076	1.1477
MONROE	1.0293	1.0325	0.0457	0.0032	1.0261	0.9704	1.0095
MONTVILLE	1.0309	1.0275	0.0398	-0.0034	1.0343	0.9851	1.0224
NAUGATUCK	1.0000	1.0328	0.0486	0.0328	0.9672	0.9138	0.9493
NEW BRITAIN	1.0771	1.0314	0.0426	-0.0457	1.1229	1.0692	1.1094
NEW CANAAN	1.0000	1.0310	0.0409	0.0310	0.9690	0.9171	0.9551
NEW FAIRFIELD	1.0000	1.0311	0.0450	0.0311	0.9689	0.9140	0.9543
NEW HAVEN	1.2461	1.0315	C.0418	-0.2146	1.4607	1.4088	1.4443
NEW INGTON	1.0231	1.0304	0.0448	0.0072	1.0159	0.9613	0.9996
NEW LONDON	1.1346	1.0299	0.0440	-0.1047	1.2394	1.1853	1.2226
NEW MILFORD	1.0000	1.0306	0.0433	0.0306	0.9694	0.9179	0.9539
NEWTOWN	1.0000	1.0285	0.0438	0.0285	0.9715	0.9188	0.9572
NORTH BRANFORD	1.0000	1.0327	0.0437	0.0327	0.9673	0.9135	0.9531
NORTH HAVEN	1.0837	1.0306	0.0436	-0.0531	1.1367	1.0815	1.1220
NORTH STONINGTON	1.0000	1.0331	0.0444	0.0331	0.9669	0.9098	0.9515

NORWALK	1.0897	1.0335	0.0452	-0.0562	1.1458	1.0923	1.1308
OLD SAYBROOK	1.0296	1.0338	0.0466	0.0042	1.0254	0.9707	1.0083
PLAINFIELD	1.0000	1.0310	0.0418	0.0310	0.9690	0.9158	0.9538
PLAINVILLE	1.0000	1.0325	0.0418	0.0325	0.9675	0.9159	0.9538
PLYMOUTH	1.0122	1.0308	0.0425	0.0187	0.9935	0.9427	0.9785
PORTLAND	1.0605	1.0333	0.0470	-0.0272	1.0876	1.0321	1.0708
PUTNAM	1.0320	1.0293	0.0436	-0.0027	1.0348	0.9823	1.0190
RIDGEFIELD	1.0000	1.0323	0.0422	0.0323	0.9677	0.9129	0.9531
ROCKY HILL	1.0676	1.0290	0.0407	-0.0386	1.1062	1.0523	1.0925
SEYMOUR	1.0000	1.0329	0.0452	0.0329	0.9671	0.9089	0.9517
SHELTON	1.0000	1.0314	0.0439	0.0314	0.9686	0.9124	0.9540
SIMSBURY	1.0000	1.0296	0.0432	0.0296	0.9704	0.9198	0.9565
SOMERS	1.0000	1.0306	0.0454	0.0306	0.9694	0.9160	0.9533
SOUTHINGTON	1.0998	1.0306	0.0449	-0.0691	1.1689	1.1153	1.1526
SOUTH WINDSOR	1.0210	1.0328	0.0469	0.0119	1.0091	0.9541	0.9909
STAFFORD	1.0000	1.0293	0.0414	0.0293	0.9707	0.9200	0.9551
STAMFORD	1.0887	1.0288	0.0407	-0.0600	1.1487	1.0981	1.1340
STONINGTON	1.0000	1.0339	0.0456	0.0339	0.9661	0.9093	0.9515
STRATFORD	1.1127	1.0313	0.0418	-0.0814	1.1941	1.1390	1.1803
SUFFIELD	1.0339	1.0299	0.0414	-0.0040	1.0380	0.9873	1.0226
THOMASTON	1.0000	1.0345	0.0423	0.0345	0.9655	0.9102	0.9527
THOMPSON	1.0000	1.0328	0.0461	0.0328	0.9672	0.9115	0.9540

TOLLAND	1.0000	1.0313	0.0419	0.0313	0.9687	0.9165	0.9530
TORRINGTON	1.0519	1.0295	0.0426	-0.0223	1.0742	1.0211	1.0598
TRUMBULL	1.0327	1.0297	0.0430	-0.0031	1.0358	0.9830	1.0203
VERNON	1.0499	1.0329	0.0432	-0.0170	1.0670	1.0131	1.0538
WALLINGFORD	1.0000	1.0334	0.0466	0.0334	0.9666	0.9096	0.9491
WATERBURY	1.0000	1.0354	0.0488	0.0354	0.9646	0.9084	0.9467
WATERFORD	1.0603	1.0338	0.0472	-0.0265	1.0869	1.0327	1.0710
WATERTOWN	1.0159	1.0323	0.0478	0.0165	0.9994	0.9465	0.9840
WESTBROOK	1.0489	1.0290	0.0425	-0.0198	1.0687	1.0173	1.0555
WEST HARTFORD	1.0000	1.0305	0.0446	0.0305	0.9695	0.9162	0.9531
WEST HAVEN	1.0000	1.0359	0.0481	0.0359	0.9641	0.9076	0.9473
WESTON	1.0000	1.0330	0.0435	0.0330	0.9670	0.9102	0.9523
WESTPORT	1.0000	1.0293	0.0447	0.0293	0.9707	0.9173	0.9571
WETHERSFIELD	1.0990	1.0352	0.0466	-0.0637	1.1627	1.1081	1.1497
WILTON	1.0000	1.0298	0.0394	0.0298	0.9702	0.9221	0.9548
WINDHAM	1.0557	1.0320	0.0445	-0.0237	1.0794	1.0265	1.0626
WINDSOR	1.0906	1.0301	0.0463	-0.0605	1.1510	1.0983	1.1346
WINDSOR LOCKS	1.0000	1.0295	0.0399	0.0295	0.9705	0.9189	0.9579
WOLCOTT	1.0475	1.0320	0.0439	-0.0156	1.0631	1.0095	1.0483
RD1	1.0589	1.0320	0.0430	-0.0269	1.0858	1.0330	1.0726
RD4	1.0518	1.0307	0.0418	-0.0211	1.0729	1.0192	1.0584
RD5	1.0375	1.0313	0.0455	-0.0062	1.0437	0.9889	1.0275

RD6	1.0000	1.0314	0.0442	0.0314	0.9686	0.9180	0.9522
RD7	1.0000	1.0294	0.0399	0.0294	0.9706	0.9206	0.9552
RD8	1.0006	1.0310	0.0407	0.0304	0.9702	0.9183	0.9556
RD9	1.0000	1.0305	0.0391	0.0305	0.9695	0.9192	0.9563
RD10	1.0460	1.0305	0.0414	-0.0155	1.0615	1.0076	1.0487
RD11	1.0797	1.0307	0.0422	-0.0489	1.1286	1.0764	1.1166
RD12	1.0871	1.0313	0.0445	-0.0558	1.1429	1.0891	1.1279
RD13	1.0515	1.0308	0.0420	-0.0207	1.0721	1.0187	1.0600
RD14	1.0000	1.0295	0.0426	0.0295	0.9705	0.9181	0.9554
RD15	1.0516	1.0288	0.0417	-0.0228	1.0744	1.0217	1.0595
RD17	1.0514	1.0292	0.0408	-0.0223	1.0737	1.0233	1.0586
RD18	1.0000	1.0330	0.0461	0.0330	0.9670	0.9123	0.9506
RD19	1.0000	1.0327	0.0429	0.0327	0.9673	0.9136	0.9527
Average	1.0314	1.0314	0.0436	0.0000	1.0315		

$\phi_i = \phi$ from the DEA model

$\bar{\bar{\phi}}_i$ = bootstrap average of ϕ

$se_B(\phi_i)$ = standard deviation of ϕ

$bias_B(\phi_i)$ = bias of ϕ

$\hat{\hat{\phi}}_i$ = bias-corrected estimate of ϕ

Table 4.8: Comparison between the two procedures (bias-corrected estimates)

	SAT	Bootstrap conditional to socioeconomic characteristics		Bootstrap using the extended DEA model with the socioeconomic characteristics	
		$\hat{\hat{\phi}}_i$	$\hat{\hat{y}}_i^f$	$\hat{\hat{\phi}}_i$	$\hat{\hat{y}}_i^f$
Bridgeport	782	1.3883	567.34	0.9690	824.91
Hartford	777	1.3224	591.39	0.9673	821.52
New Haven	789	1.3510	589.06	1.4607	1038.00
Avon	1133	.09985	1145.67	0.9699	1193.24
Glastonbury	1074	1.0105	1076.24	0.9686	1133.56
Simsbury	1140	0.9959	1158.28	0.9704	1199.44

Table 4.9: Data

School District	SAT	Mino	BA	Instrh	(T1/stu) *1000	(T2/stu) *1000	(s/stu)* 1000	(a/stu)* 1000	(pc/stu)* 1000
ANSONIA	962	26.6	17.4	947	50.0433	9.7054	4.7660	5.5459	120.4819
AVON	1133	6.3	78.9	1013	70.0219	7.6586	5.3829	7.4398	116.2791
BERLIN	987	4.5	33	964.33	61.5436	8.4134	5.3991	5.2998	161.2903
BETHEL	1051	10.3	53.5	989	61.6013	9.7331	5.8399	4.0816	178.5714
BLOOMFIELD	886	88.2	32.5	920.33	57.4819	12.9525	7.6428	4.8270	136.9863
BOLTON	1078	3.8	45.5	969.5	69.4935	6.9494	5.8893	3.5336	151.5152
BRANFORD	1027	7.6	39.1	934.33	63.1055	8.1650	5.3972	4.1517	70.4225
BRIDGEPORT	782	88.3	5.9	973	57.9023	8.1742	4.8794	4.8794	129.8701
BRISTOL	1004	13.2	21.4	954.67	60.4508	8.8597	4.5805	4.0743	90.0901
BROOKFIELD	1053	5.9	63.4	966	58.9141	8.3173	4.7362	3.8506	135.1351
CANTON	1061	3	50.8	937	63.2579	9.4229	7.3046	5.8437	86.9565
CHESHIRE	1063	5	66.7	946.67	60.4698	9.4837	6.2043	4.8748	153.8462
CLINTON	1028	9	44.3	960	60.6652	12.3725	5.0554	5.3215	140.8451
COLCHESTER	1019	3.8	39.9	920.33	52.4814	8.8916	4.9628	2.8950	93.4579
COVENTRY	1031	2.6	35.5	965	60.1156	10.6674	6.4109	4.7294	113.6364
CROMWELL	987	10.9	35	1004.6	61.9765	10.6086	5.0251	5.0251	116.2791
DANBURY	956	39.5	35.8	931	59.4988	11.0174	5.3355	5.6589	82.6446
DARIEN	1121	5.7	84.2	990	68.4160	8.8000	4.8640	5.7600	108.6957
DERBY	914	19.6	20.1	942	49.8962	7.4048	3.8062	3.4602	74.6269
EAST GRANBY	1065	4.5	49.6	976.33	65.1572	5.7862	3.7736	7.5472	204.0816

EAST HADDAM	1036	3.5	25.1	1010.3	64.0538	9.5804	3.4838	5.5424	222.2222
EAST HAMPTON	997	3.5	43.3	999	56.6459	11.1111	5.9709	5.1921	69.4444
EAST HARTFORD	973	49.4	16.9	956.67	56.9873	10.4590	4.9721	6.3254	62.8931
EAST HAVEN	908	6.3	17.5	959.33	54.6479	7.1703	4.3534	6.1460	192.3077
EAST LYME	1068	5.7	45.1	972.33	62.3066	7.7356	4.9226	4.5007	185.1852
EAST WINDSOR	963	10.9	15	970.33	57.9780	7.9092	4.1265	6.1898	156.2500
ELLINGTON	1050	3.3	41.8	987	59.4378	10.0402	8.0321	5.0201	138.8889
ENFIELD	1011	7.2	26.5	944	59.2460	8.8420	4.3387	4.2639	83.3333
FAIRFIELD	1068	7.5	60.5	958.33	62.1069	11.6616	6.4848	4.3139	151.5152
FARMINGTON	1092	11.2	65.5	998	62.9371	5.1748	4.5594	4.1958	119.0476
GLASTONBURY	1074	1.6	66.8	976.33	58.9028	9.2825	5.5049	5.5049	97.0874
GRANBY	1097	3.6	58.1	1030	58.2315	11.1693	6.3991	6.9808	128.2051
GREENWICH	1078	19.9	63.3	994	67.5158	12.7765	8.5132	7.7624	181.8182
GRISWOLD	964	4	14.6	957.67	60.4167	11.6477	5.2083	3.3144	147.0588
GROTON	1037	19.9	21.7	970	64.8217	10.3207	3.7441	5.6975	135.1351
GUILFORD	1073	4.3	59.9	947	61.3688	9.5934	5.4983	5.1546	114.9425
HAMDEN	962	27.3	45.3	1007	59.8053	9.4942	5.1221	4.9465	100.0000
HARTFORD	777	95.1	6.6	973.67	58.1102	13.5219	4.2705	4.6656	78.1250
KILLINGLY	992	6.6	15.1	983.67	62.2753	8.8264	6.2112	3.9229	125.0000
LEBANON	1036	2.1	30.5	1005.3	66.9536	8.0132	7.3510	5.2980	178.5714
LEDYARD	1045	10.6	49.7	955.33	60.6369	8.4713	6.7516	5.5414	163.9344
LITCHFIELD	1077	3.1	44.2	1013	67.0111	7.7901	4.7695	5.5644	114.9425
MADISON	1087	2.9	70.1	963	60.6780	8.1017	6.1017	4.4068	178.5714
MANCHESTER	1000	25.6	34.9	960.33	58.2667	8.9244	6.1388	4.3848	87.7193

MERIDEN	946	41.6	15.6	937	58.0197	8.6310	4.5313	3.8360	84.0336
MIDDLETOWN	961	36.9	25.8	964.67	61.1465	8.9172	7.0488	4.6709	98.0392
MILFORD	973	7.2	30.4	996.67	59.2749	9.7757	7.0532	3.8087	120.4819
MONROE	1007	5.1	58.4	956.67	54.6456	8.0486	4.8574	3.8125	111.1111
MONTVILLE	1007	8.7	25.1	957.67	61.7059	10.6353	4.6395	4.6395	196.0784
NAUGATUCK	928	8.8	22.4	925.67	49.9030	7.2096	4.0543	3.3492	109.8901
NEW BRITAIN	909	64.4	15.4	983.33	53.6089	11.4706	4.8696	4.1121	75.1880
NEW CANAAN	1129	3.9	83.5	982.67	68.2571	9.7747	5.6991	5.9642	153.8462
NEW FAIRFIELD	1048	3.7	51.4	975.33	59.9021	7.1536	5.2711	4.5181	64.9351
NEW HAVEN	789	86.6	13.8	951	55.4881	11.1839	5.1659	5.9647	142.8571
NEW INGTON	1032	10.9	40.7	974.33	61.6486	6.7374	6.9102	4.6890	133.3333
NEW LONDON	893	69.7	13.6	997.67	64.6399	15.4891	5.4348	6.4538	113.6364
NEW MILFORD	1058	6.5	45	937.67	56.8670	7.6180	4.0773	4.9356	96.1538
NEWTOWN	1065	3.2	63	958.67	60.3665	8.0070	6.1245	3.4890	111.1111
NORTH BRANFORD	1002	4.2	30.6	939.67	58.6447	7.9930	5.2129	5.2129	87.7193
NORTH HAVEN	999	8.5	40.5	957.33	61.5523	10.1888	6.6527	6.8924	111.1111
NORTH STONINGTON	1061	3.8	34.1	969	65.5098	8.6768	4.3384	4.3384	149.2537
NORWALK	943	47	28.8	934.33	60.8704	10.4543	5.0980	4.5911	102.0408
OLD SAYBROOK	1052	5.5	43.9	981.33	72.0187	11.0678	3.8971	5.4560	192.3077
PLAINFIELD	975	4.9	10.2	972.67	56.9411	11.0231	5.7182	4.4781	106.3830
PLAINVILLE	990	10.7	18.3	963	63.6256	6.1345	3.9324	5.5053	163.9344
PLYMOUTH	983	2.8	18.6	955.33	60.4368	10.4622	4.0630	5.6882	105.2632
PORTLAND	1033	7.8	31.8	993.33	69.5447	11.7739	6.2794	6.2794	158.7302
PUTNAM	986	6.4	18.4	960.67	65.8210	12.6850	4.2283	4.9331	149.2537

RIDGEFIELD	1130	4.8	78.9	957.67	63.2713	7.8599	6.7826	4.8972	114.9425
ROCKY HILL	1011	11.2	43.5	968.67	60.0000	9.0022	5.7650	3.5477	112.3596
SEYMOUR	975	5.5	27.1	998.67	49.5301	10.1801	3.1323	5.0901	114.9425
SHELTON	982	7.9	38.5	971.67	53.7296	9.9394	4.0704	4.7331	73.5294
SIMSBURY	1140	5.7	81.6	927	61.2692	8.3528	6.6262	4.6664	138.8889
SOMERS	1017	2	41.8	1041	60.9396	8.0537	3.6913	6.7114	232.5581
SOUTHINGTON	993	5	38.7	972.67	61.9469	10.4021	6.0860	4.5024	112.3596
SOUTH WINDSOR	1051	8.8	51.2	963	57.4581	8.2740	5.7458	4.8265	95.2381
STAFFORD	1080	4.2	13.9	1013.3	53.5679	11.6947	7.4331	4.4599	76.3359
STAMFORD	950	54.6	32	932.33	64.5780	10.2139	5.9934	3.8042	136.9863
STONINGTON	1034	3.5	34.6	926.33	65.6696	12.3214	4.0179	5.8036	135.1351
STRATFORD	948	24.4	31.6	984	57.7191	7.8600	7.8600	5.9321	166.6667
SUFFIELD	1026	3.7	48.6	1058.6	61.2328	8.6093	4.5848	3.3622	107.5269
THOMASTON	937	2.5	21.2	956.67	58.4133	7.8466	3.0514	6.1029	54.9451
THOMPSON	1060	1.7	22.5	995.33	54.0816	9.4558	2.9932	5.4422	149.2537
TOLLAND	1089	3.4	49.8	949.33	58.7905	9.1145	4.4924	3.4557	104.1667
TORRINGTON	970	8.6	25.9	946	53.7087	8.6295	5.3216	4.3148	188.6792
TRUMBULL	1029	7.8	66.8	983.33	56.3535	8.0186	7.5907	4.0930	114.9425
VERNON	1036	14	38.7	967.67	63.5864	11.1270	7.0678	6.6858	100.0000
WALLINGFORD	968	7.8	33.1	940	60.8311	7.2207	4.1261	3.8314	71.9424
WATERBURY	856	59.8	10.3	924	59.5809	11.8337	2.9340	4.6245	49.5050
WATERFORD	1033	7.8	29.4	1014.6	68.4230	12.2318	8.4682	5.2691	178.5714
WATERTOWN	1000	3.4	37.9	962.33	56.3396	9.0361	3.5858	4.0161	135.1351
WESTBROOK	1024	5.2	60.2	997.33	71.3939	7.8788	3.6364	6.0606	181.8182

WEST HARTFORD	1070	21.9	65.2	944.67	60.8588	9.5199	6.4410	5.0725	72.9927
WEST HAVEN	935	35.2	18.7	969	53.1609	8.6344	3.4209	3.0104	60.2410
WESTON	1110	4.4	93.4	981	68.1952	9.7002	4.7031	5.8789	153.8462
WESTPORT	1139	5.8	78.9	1027.6	66.6842	16.2497	7.6376	6.8475	188.6792
WETHERSFIELD	1012	10	50.1	966.33	60.6796	10.6149	7.0550	6.4725	114.9425
WILTON	1142	5	83.9	1022	60.9841	7.6825	7.0794	4.4444	200.0000
WINDHAM	992	44.9	17.4	973	67.3244	14.3064	7.3709	4.9333	121.9512
WINDSOR	989	42.4	47.3	947.67	57.8017	8.5964	7.7905	4.2534	129.8701
WINDSOR LOCKS	1008	9.6	22.3	956	52.3654	7.1234	5.0027	4.8940	135.1351
WOLCOTT	942	3.8	25.5	937	56.7520	8.5643	5.4016	4.3355	87.7193
RD1	1005	3.8	36.7	974	88.4462	8.9641	5.9761	13.9442	89.2857
RD4	1053	2.9	46.2	966.5	78.3820	7.9576	6.8966	9.2838	285.7143
RD5	1080	8.5	72.8	933.5	79.2283	7.9281	11.6279	5.2854	156.2500
RD6	1116	2.4	45.3	971.5	73.0769	8.1731	6.3462	8.6538	161.2903
RD7	1089	1.4	46.6	959	70.4989	7.5922	7.7007	6.5076	196.0784
RD8	1096	2.2	53.6	967	79.2088	9.0909	6.3131	6.1448	158.7302
RD9	1090	2.9	70.6	995	69.3215	6.4897	8.2596	10.3245	131.5789
RD10	1049	2.8	49	1031.6	62.8706	9.4542	7.0907	6.6609	128.2051
RD11	1014	4.2	26.8	1038	91.0864	13.9276	8.3565	13.9276	312.5000
RD12	1010	3.1	48.3	957.33	67.3893	10.5561	5.0895	6.5975	204.0816
RD13	1039	3.5	47.3	1016.6	64.9671	9.1009	4.2215	6.0307	129.8701
RD14	1002	4.4	44.3	970.67	61.8209	8.7127	3.7200	3.4263	169.4915
RD15	1028	3.7	51.1	954.33	56.4030	10.3574	5.9949	5.3476	128.2051
RD17	1049	3.1	60.1	1039	69.2010	9.2978	5.2300	5.3269	175.4386

RD18	1117	2.9	56.3	996.33	69.8276	11.7816	5.3161	6.4655	136.9863
RD19	1119	8.4	56.6	1032	78.3439	6.3694	7.4310	5.3079	208.3333

Chapter 5: Dissertation Summary

5.1 Introduction

The dissertation consists of three essays that are developed independently. They all contribute to estimation of statistical production frontiers and technical efficiency using alternative techniques. The first essay relates to econometric estimation of production frontiers from panel data and modeling of technical efficiency and technical change. The second essay develops a methodology for estimation of a statistical production frontier using mathematical programming and bootstrapping techniques. Finally, the third essay estimates a non-parametric frontier using Data Envelopment Analysis and bootstrapping. Each essay includes a methodological extension with an empirical application.

The first chapter provides a brief literature review and places the three essays in perspective. Chapters 2 to 4 present the three essays, which are summarized in the following sections of this chapter.

5.2 Frontier Production Function Models with Autoregressively Time-Varying Efficiency

Traditional econometric estimation techniques fail to measure a production frontier, because they allow the observed output bundle produced by a given set of inputs to be greater than the estimated maximal producible output. The composed error frontier was first introduced by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977). The composed error is the sum of a two-sided error term that represents the random shocks and a one-sided error term that represents technical inefficiency. Although this procedure has been extended to panel data, researchers usually model the technical efficiency as an explicit function of time. This does not allow us to distinguish between technical change and efficiency change. The first essay in this dissertation constructs a model specifying technical efficiency change through firm-specific intercepts that evolve over time as a first order auto-regressive process. Besides allowing efficiency in one period to be influenced by past levels of efficiency, this approach separates efficiency from technical change.

For the empirical application the first essay uses a panel data set of 12 states from the Indian Manufacturing sector for the period 1954-1983.

5.3 Mathematical Programming Estimation of a Parametric Production Frontier

A major drawback of econometric estimation of composed error frontiers is that explicit assumptions are required about the probability distribution of the error terms. Such assumptions are arbitrary and alternative assumptions can lead to different conclusions about the technical efficiency of a firm. The second essay returns to the Aigner and Chu (1968) approach of modeling a deterministic frontier using mathematical programming, but estimates a parametric stochastic production function with a composed error term instead of a one-sided error term. The individual levels of technical efficiency can be estimated without any restrictive assumption about the statistical distribution of the error terms. As with any mathematical programming approach, we get only point estimates for the parameters of interest. In order to overcome this problem, bootstrapping techniques are employed and confidence intervals for parameters are constructed.

The methodology developed in the second essay uses data from the US Census of Manufacturing, 1992.

5.4 A Bootstrap Procedure to Capture Unit Specific Effects In Data Envelopment Analysis

A drawback of any parametric frontier estimation is the subjective choice of the functional form of the frontier. Data Envelopment Analysis, which was first introduced by Charnes, Cooper and Rhodes (1978, 1981) and is based on the assumptions of monotonicity, convexity and free disposability of inputs and outputs, avoids any functional specification. But, it is a mathematical programming approach and can only provide point estimates of the relative technical efficiency of a firm. One proposed solution to this problem is to use bootstrapping. Simar (1992, 1996) and Simar and Wilson (1997a, 1997b) set the foundation for consistent use of bootstrapping. A problem with this approach is that it assumes that all the firms in the sample have the same probability of getting an observed technical efficiency level, while the firm's relative efficiency position might be systematically influenced by unit specific factors out of the firm's control. The third essay in this dissertation develops a bootstrap procedure that generates the distribution of efficiency for each firm, conditional on unit specific factors. Additionally, we construct confidence intervals not for the technical efficiency scores but for the expected maximum producible output.

The essay in the fourth chapter includes an application with Connecticut school data for the academic year of 1995-1996.

5.4 Directions for Future Research

A potential extension of the frontier production function model with autoregressively time-varying technical efficiency is to specify a one-sided distribution for the technical efficiency term, for the half-normal distribution. Also, it will be very interesting to estimate the model using Kalman filter and to compare the results from the other two model specifications.

Similar to the bootstrap procedure developed in Chapter 4, a bootstrap procedure that captures unit specific effects can be also developed for the Free Disposal Hull Analysis (FDH) framework, which is an alternative to Data Envelopment Analysis (DEA). FDH avoids the convexity assumption for the production technology and thus, is less restrictive.

REFERENCES

Afriat, S. (1972). "Efficiency Estimation of Production Functions." *International Economic Review*, 13, 3, 568-598.

Ahn, S. C., Y. H. Lee, and P. Schmidt. (1995). "GMM estimation of a panel data regression model." unpublished manuscript, Arizona State University.

Aigner, D. J., and S. F. Chu (1968). "On Estimating the Industry Production Function." *American Economic Review 58*, 826-839.

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

Banker, Rajiv D. (October 1993). "Maximum Likelihood, Consistency and Data Envelopment Analysis: A Statistical Foundation." *Management Science*, vol.39, no. 10, 1265-1273.

Banker, Rajiv D., A. Charnes, and W. W. Cooper. (September 1984). "Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis." *Management Science*, vol. 30, no. 9, 1078-1092.

Banker, R. D., and A. Maindiratta. (1992). "Maximum Likelihood Estimation of Monotone and Concave Production Frontiers." *Journal of Productivity Analysis*, 401-415.

Bauer, Paul W. (1990) "Recent Developments in the Econometric Estimation of Frontiers." *Journal of Econometrics 46*, 39-56.

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

Battese, George E., and Tim J. Coelli. (1992). "Frontier Production Functions, Technical Efficiency and Panel Data: An Application to Paddy Farmers in India." *Journal of Productivity Analysis 3*, 153-169.

Battese, George E., and Tim J. Coelli. (1993). "A Generalized Frontier Production Function Incorporating a Model for Inefficiency Effects." *mimeo*, Department of Econometrics, University of New England, Armidale, Australia.

Battese, George E., and Tim J. Coelli. (1995). "A Model for Technical Inefficiency Effects in a Stochastic Production Function for Panel Data." *Empirical Economics* 20, 325-332.

Bauer, Paul W. (1990) "Recent Developments in the Econometric Estimation of Frontiers." *Journal of Econometrics* 46, 39-56.

Bera, Anil K., and Subhash C. Sharma. (1996). "Estimating Production Uncertainty in Stochastic Frontier Production Function Models." working paper 96-0124, University of Illinois at Urbana-Champaign.

Beran, Rudlof, and Gilles R. Ducharme. (1991). *Asymptotic Theory for Bootstrap Methods in Statistics.* Centre de Recherches Mathématiques, Université de Montréal, Québec, Canada.

Bureau of Evaluation and Student Assessment. (1996). *Strategic School Profiles.* Connecticut State of Education, Hartford, Connecticut.

Charnes, A. W. W. Cooper, Z. M. Huang, and D. B. Sun. (1990). "Polyhedral Cone-Ratio DEA Models with an Illustrative Application to Large Commercial Banks. " *Journal Of Econometrics 46,* 73-91.

Charnes, A. W. W. Cooper, and E. Rhodes. (1978). "Measuring the Efficiency of Decision Making Units." *European Journal of Operational Research*, 429-444.

Charnes, A. W. W. Cooper, and E. Rhodes. (1981). "Evaluating Program and Managerial Efficiency: An Application of Data Envelopment Analysis to Program Follow Through." *Management Science*, vol. 27, no. 6, 668-697.

Coelli, Tim J. (1995). "Estimators and Hypothesis Tests for a Stochastic Frontier Function: A Monte Carlo Analysis." *Journal of Productivity Analysis*, vol.6, no. 3, 247-268.

Cooley, Thomas, F, and Edward C. Prescott. (1973a). "An Adaptive Regression Model." *International Economic Review*, 2, 463-474.

Cooley, Thomas, F, and Edward C. Prescott. (1973b). "Varying Parameter Regression. A Theory and Some Applications." *Annals of Economic and Social Measurement*, 2, 463-474.

Cooley, Thomas, F, and Edward C. Prescott. (1976). "Estimation in the Presence of Stochastic Parameter Variation." *Econometrica*, 44, 167-184.

Cornwell, Christopher, Peter Schmidt, and Robin C. Sickles. (1990). "Production Frontiers with Cross Sectional and Time-Series Variation in Efficiency Levels." *Journal of Econometrics*, 46, 185-200.

Debreu, G. (1951). "The Coefficient of Resource Utilization." *Econometrica*, 19, 3, 273-292.

Deprins, Dominique, Lèopold Simar, and Henry Tulkens. (1984). "Measuring Labor-Efficiency in Post Offices." *The Performanve of Public Enterprises: Concepts and Measurement.*

Efron, Bradley. (1979). "Bootstrap methods: Another Look at the Jackknife." *Annals of Statistics*, vol. 7, 1-26.

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

Farrell, M. (1957). "The measurement of Productive Efficiency." *Journal of the Royal Statistical Society* 120, pt. 3, 253-281.

Fuss, Melvyn A., and Leonard Waverman. (1992). *Cost of Productivity in Automobile Production. The Challenge of Japanese Efficiency.* New York: Cambridge University Press.

Graybill, Franklin, A. (1969). *Introduction to Matrices with Applications in Statistics.* Belmont, California: Wadsworth Publishing Company, Inc.

Greene, William. (1980a). "Maximum Likelihood Estimation of Econometric Frontier Functions." *Journal of Econometrics 13,* 27-56.

Greene, William. (1980b). "On the Estimation of a Flexible Frontier Production Model." *Journal of Econometrics 13,* 101-115.

Greene, William. (1990). "A Gamma Distributed Stochastic Frontier Model." *Journal of Econometrics 46,* 141-163.

Greene, William. (1993). "Chapter 2: Frontier Production Functions." in *The Measurement of Productive Efficiency,* (eds.), Fried, H., K. Lovell, and S. Schmidt. New York: Oxford University Press.

Greene, William. (1995). "Frontier Production Functions." in *Handbook of Applied Econometrics, Vol. II-Microeconomics,* (eds.), H. Pesaran and P. Schmidt London: Basil Blackwell.

Greene, William. (1997). *Econometric Analysis.* 3rd edition, New Jersey: Prentice Hall.

Hamilton, James D. (1994). *Time Series Analysis.* New Jersey: Princeton University Press.

Hesmati, Almas, and Kumbhakar, Subal C. (1994). "Farm Heterogeneity and Technical Efficiency: Some Results from Swedish Dairy." *Journal of Productivity Analysis*, 5, 45-61.

Hjalmarsson, Lennart, Subal C. Kumbhakar, and Almas Heshmati. (1995). "DEA, DFA, and SFA: A Comparison." *Journal of Productivity Analysis,* 7, 303-327.

Horrace, W. C., and P. Schmidt. (1995). "Multiple Comparisons with the Best, with Applications to the Efficiency Measurement Problem." unpublished manuscript, Department of Economics, Michigan State University.

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

Huang, Cliff J., and Jin-Tan Liu. (1994). "Estimation of a Non-Neutral Stochastic Frontier Production Function." *Journal of Productivity Analysis 5,* 171-180.

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

Judge, George G., W. E. Griffiths, R. Carter Hill, Helmut Lütkepohl, and Tsoung-Chao Lee. (1988). *Introduction to the Theory and Practice in Econometrics.* 2nd edition, New York: John Wiley and Sons.

Judge, George G., W. E. Griffiths, R. Carter Hill, Helmut Lütkepohl, and Tsoung-Chao Lee. (1985). *The Theory and Practice in Econometrics.* 2nd edition, New York: John Wiley and Sons.

Kumbhakar, Subal C. (1990). "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency." *Journal of Econometrics,* 46, 201-211.

Kumbhakar, Subal C. (1991). "Estimation of Technical Inefficiency in Panel Data Models with Firm- and Time- Specific Effects." *Economic Letters*, 36, 43-48.

Kumbhakar, Subal C., and Almas Hesmati. (August 1995). "Efficiency Measurement in Swedish Dairy Farms: An Application of Rotating Panel Data." *American Journal of Agricultural Economics,* 77, 660-674.

Land, Kenneth C., Knox C. A. Lovell, and Sten Thore. (1988). " Analysis." Paper presented at the National Science Foundation Conference, Chape; Hill, Inc., 1988.

Land, Kenneth C., Knox C. A. Lovell, and Sten Thore. (1993). "Chance-constrained Data Envelopment Analysis." *Managerial and Decision Economics,* vol.14, 541-554.

Lee, Young Hoon, and Peter Schmidt. (1993). " A production Frontier Model with Flexible Temporal Variation In Technical Efficiency." in *The Measurement of Productive Efficiency. Techniques and Applications.* eds. Harold O. Fried, C. A. Knox Lovell, and Shelton S. Schmidt, New York: Oxford University Press.

Maddala, G. S. (1982). *Econometrics*, 4th edition, Tokyo: McGraw-Hill Inc.

McCarty, Therese A., and Suthathip Yaisawarng. (1993). "Technical efficiency in New Jersey School Districts." in the *Measurement of Productive Efficiency. Techniques and Applications.* eds. Harold O. Fried, C. A. Knox Lovell, and Shelton S. Schmidt, New York: Oxford University Press.

Meeusen, Wim, and Julien van den Broeck. (June 1977). "Efficiency Estimation from Cobb-Douglas Production Functions with Decomposed Error." *International Economic Review*, vol. 18, no. 2, 435-444.

Olesen, O. B., and N. C. Petersen (1989). " Chance-Constrained Efficiency Evaluation." *Publications from Department of Management*, no. 9/1989. Odense University, Denmark.

Olesen, O. B., and N. C. Petersen (1995). "Incorporating Quality into Data Envelopment Analysis: A Stochastic Dominance Approach." *International Journal of Production Economics* 39, 117-135.

Olson Jerome A., Peter Schmidt, and Donald M. Waldman. (1980). " A Monte Carlo Study of Estimators of Stochastic Frontier Production Functions." *Journal of Econometrics 13*, pp. 67-82.

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

van de Panne, C. (1975). *Methods for Linear and Quadratic Programming*. Amsterdam: North-Holland Publishing Company.

Ray, Subhash C. (December 1991). "Resource-use Efficiency in Public Schools: A Study of Connecticut Data." *Management Science*, vol. 37, no. 12, 1620-1628.

Reifschneider, David, and Rodney Stevenson. (August 1991). "Systematic Departures from the Frontier: A Framework for the Analysis of Firm Inefficiency." *International Economic Review*, vol. 32, no. 3, 715-723.

Richmond, J. (June 1974). "Estimating the Efficiency of Production." *International Economic Review*, vol.15, no. 2, 515-521.

Rosenberg, Barr. (1973). "Random Coefficients Models. The Analysis of a Cross Section of Time Series by Stochastically Convergent Parameter Regression." Annals of Economic and Social Measurement, 2/4, pp. 399-428.

SAS Institute Inc. (1989). *SAS/IMLSoftware: Usage and Reference, Version 6, First Edition*, Cary, NC: SAS Institute Inc.

SAS Institute Inc. (1995). *SAS/IMLSoftware: Changes and Enhancements through Release 6.11*, Cary, NC: SAS Institute Inc.

Sarath, Bharat, and Ajay Maindiratta. (1997). "On the Consistency of Maximum Likelihood Estimation of Monotone and Concave Production Frontiers." *Journal of Productivity Analysis*, 8, 239-246.

Schmidt, Peter. (1976). "On the Statistical Estimation of Parametric Frontier Production Functions." *Review of Economics and Statistics*, 238-239.

Schmidt, Peter and Robin C. Sickles. (October 1984). "Production Frontiers and Panel Data." *Journal of Business and Economic Statistics*, vol. 2, no. 4, 367-374.

Silverman, B. W. (1986). *Density Estimation of Statistics and Data Analysis*. London: Chapman and Hall.

Simar, Léopold. (June 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*, vol. 3, no 1/2, 167-203.

Simar, Léopold. (1996). "Aspects of Statistical Analysis in DEA-Type Frontier Models." *Journal of Productivity Analysis*, 7, 177-185.

Simar, Léopold, and Paul W. Wilson. (1997a). "Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Models. " *forthcoming to Management Sciences*.

Simar, Léopold, and Paul W. Wilson. (1997b). "Nonparametric Tests of Returns to Scale." presented at the Fifth European Productivity Workshop on Efficiency and Productivity Analysis, Copenhagen.

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

Thompson, Russell G., P. S. Dharmapala, and Robert M. Thrall. (1995). "Linked-cone DEA Profit Ratios and Technical Efficiency with Application to Illinois Coal Mines." *International Journal of Production Economics* 39, 99-115.

Thompson, Russell G., Larry N. Langemeier, Chih-Tah Lee, Euntaik Lee, and Robert M. Thrall. (1990). "The Role of Multiplier Bounds in Efficiency Analysis with Application to Kansas." *Journal of Econometrics* 46, 93-108.

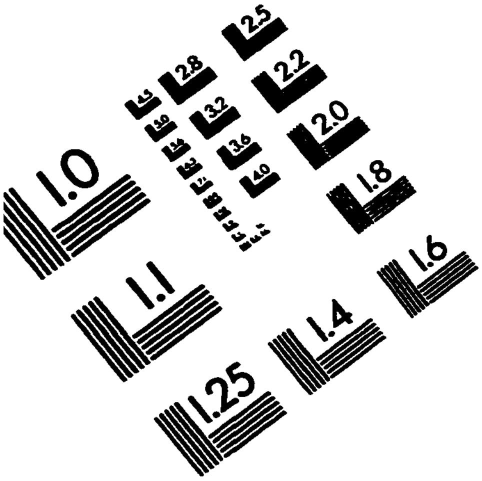
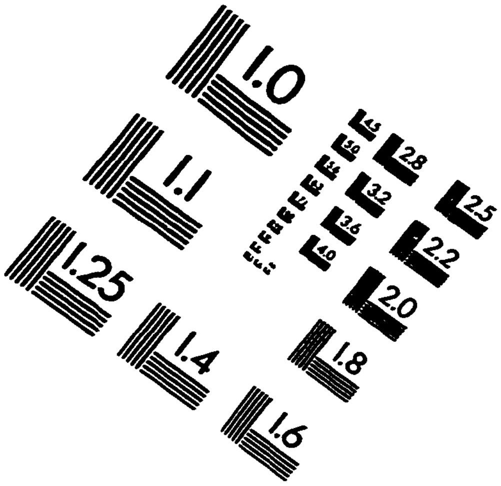
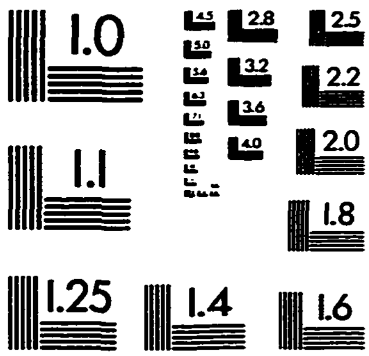
Timmer, P. (1971). "Using a Probabilistic Frontier Production Function to Measure Technical Efficiency." *Journal of Political Economy*, 79, 776-794.

Varian, Hal R. (1985). "Non-Parametric Analysis of Optimizing Behavior with Measurement Error." *Journal of Econometrics* 30, 445-458.

Yancey, T. A., G. G. Judge, and M. E. Bock. (1981). "Testing Multiple Equality and Inequality Hypotheses in Economics." *Economics Letters*, 4, 249-255.

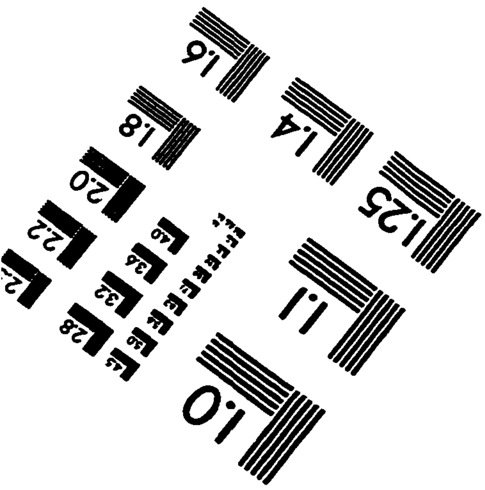
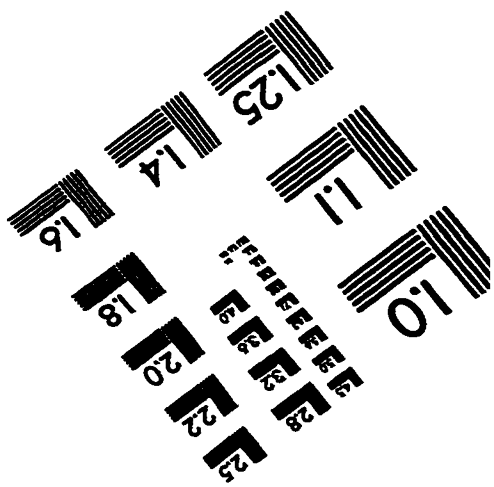
Yancey, T. A., R. Bohrer, and G. G. Judge. (1982). "Power Function Comparisons in Hypothesis Testing." *Economics Letters*, 5, 161-167.

IMAGE EVALUATION TEST TARGET (QA-3)

150mm

6"

APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989