UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Nonparametric Measurement of Technical Efficiency Using Panel Data

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Economics

by

Daniel Joseph Henderson

June, 2003

Dissertation Committee:

    Professor R. Robert Russell, Co-Chairperson
    Professor Aman Ullah, Co-Chairperson
    Professor Jang-Ting Guo
    Professor Tae-Hwy Lee

The Dissertation of Daniel Joseph Henderson is approved:

_____

_____

_____
Committee Co-Chairperson

_____
Committee Co-Chairperson

University of California, Riverside

Pariyaprasert, Albert Pasaoa, Gurleen Popli and Heather Tiereny. I would also like to thank the department itself for its financial and professional assistance.

Finally, I want to thank my family for the tremendous amount of support. I would like to thank my wife for being there through all of the struggles (both emotional and financial). You already know that you are the reason that I am able to accomplish this work. Remember that although sometimes my mind is in the clouds, you are always in my heart. Again, I cannot find the words to express the love and respect that I have for you. I would also like to give thanks to my beautiful sisters, I am very proud of each of you. I know that you will succeed in whatever directions life takes you. To the person who taught me to be a man, my father; I can only hope to be half the man that you are (we'll overlook the fact that you did not teach me how to switch hit). Lastly (and I meant to write this last because it is the most difficult), I want thank those whom I can no longer thank in person. I want to thank my grandmother and grandfather (Henderson). Without your struggles, sacrifices and courage, none of us could have achieved any of this. The most emotional and true thanks I have are for you mother. You had faith in me when others did not. You cared for me at whatever the cost. You did what was best for my career and goals regardless of how that would affect you. I think about you everyday. I think about how you pretended to be okay so that I could study for my comprehensive exams, knowing you only had days to live. I can only imagine the pain that you must have felt, but I realize that you would make the same decision, had you been placed in that situation a hundred times. That is why I have your picture next to my desk so that you can watch over me while I work. You have been a tremendous inspiration. I can

remember our last conversation; when you said that you were going to be there at my graduation. I now know that you meant spiritually and not physically. Mother (Susan Marie Henderson), you have always and will always be with me.

DEDICATION

to all my family and friends

ABSTRACT OF THE DISSERTATION

The Nonparametric Measurement of Technical Efficiency Using Panel Data

by

Daniel Joseph Henderson

Doctor of Philosophy, Graduate Program in Economics
University of California, Riverside, June, 2003
Professor R. Robert Russell, Co-Chairperson
Professor Aman Ullah, Co-Chairperson

In recent years, much attention has been made to the developments and extensions of technical efficiency measurement using panel data. The two major strands of research are deterministic (where all observations lie on one side of the frontier) and stochastic (where observations lie on both sides of the frontier) frontier measures. In this dissertation, the first chapter focuses on summarizing the existing research.

Although deterministic methods are often chosen over stochastic (parametric) measures because they do not assume a functional form, many econometricians disapprove of them because (in general) they lack statistical properties. Nonparametric kernel methods have the potential to satisfy both of these wants. The second chapter of the dissertation considers the problem of improving the estimation of the general one-way random effects (random effects procedures have been shown to perform well in this literature with most economic panel data) error component model (this development is a necessary first step which is needed in order to develop a nonparametric model for technical efficiency). We

propose the first fully nonparametric random effects kernel estimator and define its structure. In addition, we prove that the estimators are consistent and asymptotically normal. Our Monte Carlo exercises show that the estimator performs almost as well as the parametric estimator (in mean squared error) when the true technology is linear, but drastically outperform the parametric model when the technology becomes nonlinear.

These advances allow us, in the final chapter, to develop a nonparametric kernel procedure in order to estimate production (or cost) frontiers as well as estimate technical efficiency. We propose the estimator, define its structure and expose its properties. Further, we find through Monte Carlo exercises that when the technology is nonlinear that the nonparametric estimates of technical efficiency have higher correlation with the true inefficiencies. We then apply these techniques in order to estimate technical efficiency levels of 17 OECD economies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
THE MEASUREMENT OF TECHNICAL EFFICIENCY USING PANEL DATA

1 Introduction

Output-oriented technical efficiency refers to a firm's ability to obtain maximum output from a given amount of inputs.[1] Formally, the level of technical efficiency is measured by the distance a particular firm is from the production frontier. Thus, a firm that sits on the production frontier is said to be technically efficient. This concept is important to firms because their profits depend highly upon their value of technical efficiency. Two firms, that have identical technologies and inputs but different levels of technical efficiency, will have different levels of output. This will create a higher revenue for one firm although both have the same costs, obviously generating a larger profit for the more efficient firm.

The main reasons for examining technical efficiency as opposed to another type of efficiency are expressed by Kumbhakar and Lovell (2000). They state that technical efficiency is a purely physical notion that can be measured without recourse of price information and having to impose a behavioral objective on producers. It is well known that price data is often difficult to find and/or flawed. For this reason alone, one might decide to focus on technical efficiency. On the other hand, cost, revenue and profit efficiency are economic concepts whose measurement requires both price information and the imposition of an appropriate behavioral objective on producers.

---

[1]This paper will examine output-oriented technical efficiency as opposed to input-oriented technical efficiency which refers to a firm's ability to minimize inputs from a given amount of output. Note that Atkinson and Cornwell (1993) show that these two values are only the same under the assumption of constant returns to scale. Further note, unless otherwise stated, that output-oriented technical efficiency will be referred to as technical efficiency for the remainder of the paper.

In addition, measuring output based technical efficiency seems to be more relevant in real life scenarios. A firm could more easily attempt to increase output with a given amount of inputs rather than decrease inputs to produce a given amount of output. In many cases, inputs lack liquidity or are costly to eliminate (e.g. unemployment benefits).

This chapter will survey the developments of technical efficiency measurement using panel data. It will primarily focus on both deterministic (where all observations lie on one side of the frontier) and stochastic (where observations lie on both sides of the frontier) production frontiers as well as their extensions. A deterministic production frontier, also known as a non-stochastic or nonparametric production frontier, is constructed by a linear programming technique. This is the method most often used by those in microeconomics, management science or operational research.[2] In contrast, the stochastic production frontier is constructed using econometric estimation. This method is most often used by econometricians.

Attempts to measure production frontiers empirically begin with Farrell (1957). In his pioneering work, Farrell provides a measure of productive efficiency as well as a definition for a production frontier. To obtain the production frontier, he uses a linear programming method. Subsequently, this method forms the basis of the Data Envelopment Analysis (DEA) method by Charnes, Cooper and Rhodes (1978).[3]

The alternative to the deterministic frontier, the stochastic frontier, originated by

---

[2]For a management science or an operational research approach, one should consult papers such as Charnes, Cooper, Lewin and Seiford (1994) for cross-sectional data and Tulkens and Vanden Eeckaut (1995) for panel data.

[3]It may be more appropiate to call this method deterministic frontier analysis (DFA) since it was first developed by Farrell (1957), but due to the popularity of the term DEA, it will be used in the place of DFA throughout the remainder of the paper.

Meeusen and van den Broeck (1977) and Aigner, Lovell and Schmidt (1977), uses econometric methods to estimate the frontier. It is important to note that these first models concentrate on cross-sectional data and estimate technical efficiency using maximum likelihood estimation (MLE). Schmidt and Sickles (1984) point out three main difficulties concerning maximum likelihood (ML) methods and consistency of estimates from using cross-sectional data. First, technical efficiency of a particular firm can be estimated, but not consistently. Next, distributional assumptions are required about technical efficiency in order to estimate the model and separate technical efficiency from statistical noise. Finally, it may be incorrect to assume that efficiency is independent of the regressors. Each of these difficulties are potentially avoidable if a "satisfactory" panel data set is available.[4]

The potential gains from using panel data to measure technical efficiency appear to be quite large. A panel obviously contains more information about a particular firm than does a cross-section of the data. Moreover, Schmidt and Sickles (1984) suggest that panel data will enable one to relax some of the strong assumptions that are related to efficiency measurement in the cross-sectional framework. Pitt and Lee (1981) extend the cross-sectional Maximum Likelihood (ML) technique to analyze a panel data set. Branching from this, Schmidt and Sickles (1984) apply random and fixed effects procedures on a panel toward the estimation of a stochastic production frontier in order to estimate time-invariant technical efficiency. Many papers are being published expanding from these works, some of which will be discussed in the third section.

The remainder of the chapter is roughly outlined in the following manner. Section 2

---

[4]A "satisfactory" panel data set is one in which the number of cross-sectional units (or firms) is large and the time length of the panel tends toward infinity. However, this is not always the case. For the benefits and limitations of standard panel data and error correction models as well as econometric estimation, one should consult Hsiao (1986) or Baltagi (1995).

is devoted to studying deterministic frontiers, while section 3 will examine stochastic frontiers and their recent developments. Possible suggestions for future research will compose the fourth section and the conclusion makes up the final section.

2 Deterministic Frontiers

Farrell's deterministic approach to estimating production functions spawned a number of similar formulations. Each of these having slightly different assumptions. The two main approaches are DEA, and the free disposable hull (FDH). DEA can be further broken into three separate subsets according to the assumption on the returns to scale. For each of these approaches, output-possibility sets and efficiency degrees in output for panel data will be given.

2.1 Data Envelopment Analysis

Efficiency is measured, in the DEA literature, by the distance between a production plan and the production frontier. Let

$$B^t = \{(u^t, x^t)\} \cup \{(0^M, 0^N)\}, \tag{1}$$

such that (u^t, x^t) is feasible and where u are the $m = 1, 2, ..., M$ outputs, x are the $n = 1, 2, .., N$ inputs and t are the $t = 1, 2, ..., T$ time periods, represent a set of actually observed production plans. A production plan in the interior of the production frontier is considered inefficient, while a point on the frontier is considered efficient.[5]

To construct the production frontier, three assumptions must be established. First,

---

[5]Some in the literature refer to the point on the empirical production frontier to be a "best practice" point. This is because it is often argued that no realistic firm can possibly be 100% efficient. Note that although this notion of efficiency is implausible and should be replaced with "best practice", the term efficiency will be used in its place throughout the paper as to follow the trend of the literature.

every observed production plan belongs to the production set. This makes the DEA analysis a deterministic one. Second, any unobserved production plan that is weakly dominated by another production plan is also part of the production set. This assumption allows for free disposability. The third assumption concerns the issue of combinations of production plans and has several different forms. The form of the third assumption will determine the assumed returns to scale. This will often affect the calculated levels of efficiency.

If the returns to scale are assumed to be constant then the third assumption would read: every unobserved production plan that is a linear combination of production plans is itself part of the output-possibility set. In conjunction with the first two assumptions, the panel data output-possibility set is given by

$$P_{CRS}^t(x^t|M^t, N^t) = \{u^t \in R_+^M|u^t \le z^t M^t \wedge z^t N^t \le x^t, z^t \in R_+^N\}, \tag{2}$$

where z^t are the slack variables.

Similarly, if the returns to scale are assumed to be non-increasing, then the third assumption states; every unobserved production plan which is a convex combination of production plans, including the origin, also belongs to the output-possibility set. The new output-possibility set is defined as

$$P_{NRS}^t(x^t|M^t, N^t) = \{u^t \in R_+^M|u^t \le z^t M^t \wedge z^t N^t \le x^t, z^t \in R_+^N \wedge \sum_n z^t \le 1\}. \tag{3}$$

The final variation of the returns to scale assumption is that of variable returns. The third assumption is similar to the previous except that the origin is no longer included in the output-possibility set. The output-possibility set becomes

$$P_{VRS}^t(x^t|M^t, N^t) = \{u^t \in R_+^M | u^t \leq z^t M^t \wedge z^t N^t \leq x^t, z^t \in R_+^N \wedge \sum_n z^t = 1\}. \quad (4)$$

Tulken and Vanden Eeckaut (1995) describe this set as the convex free disposal hull of the data.[6]

Now that the essential groundwork has been laid, the numerical measurement of efficiency can be obtained. In order to determine an output efficiency degree relative to a DEA CRS output-possibility set, the optimal λ^{t*} must be obtained from the following problem

$$\min_{z^t, \lambda^t} \lambda^t$$

$$\text{subject to} \quad \frac{u^t}{\lambda^t} \leq z^t M^t \quad (5)$$

$$z^t N^t \leq x^t$$

$$z^t \geq 0.$$

Note that $\lambda^{t*} \leq 1$, but that it is feasible that $\lambda^{t*} = 1$. When the equality holds the firm is said to be efficient. The efficiency degree in output relative to the P_{CRS}^t output-possibility set is

$$E_{CRS}^t(u^t, x^t|M^t, N^t) = \lambda_{CRS}^{t*}. \quad (6)$$

If

$$\sum_n z^t \leq 1,$$

---

[6]In comparing the three methods, it should be noted that $P_{CRS}^t \supseteq P_{NRS}^t \supseteq P_{VRS}^t$. These are obviously formed by the differences in the third assumption.

is added to the bottom of the problem, then the efficiency degree in output relative to the P_{NRS}^t output-possibility set is

$$E_{NRS}^t(u^t, x^t | M^t, N^t) = \lambda_{NRS}^{t*}.$$ (7)

Finally, to find the efficiency degree in output relative to the P_{VRS}^t output-possibility set, the final line of the maximization problem is replaced by

$$\sum_n z^t = 1,$$

and the efficiency degree is

$$E_{VRS}^t(u^t, x^t | M^t, N^t) = \lambda_{VRS}^{t*}.$$ (8)

2.2   Free Disposable Hull

Not everyone is comfortable with the third assumption in the DEA framework. Convexity and/or the returns to scale assumptions seem too restrictive to some authors. In contrast, the FDH has nearly identical modeling features and properties, except that it omits a third assumption. The FDH output-possibility set for panel data is

$$P_{FDH}^t(x^t | M^t, N^t) = \{u^t \in R_+^M | u^t \le z^t M^t \wedge z^t M^t \le x^t, \sum_n z^t = 1 \wedge z^t \in \{0, 1\}\}.$$ (9)

Similar to DEA, the FDH production set is deterministic and as indicated by its name, also allows for free disposability.

When determining output efficiency from the output-possibility set, the processes is nearly identical to the DEA measures except that this additional line is added to the previous maximization problem

$$z^t \in \{0, 1\}.$$

In addition, the measure of output efficiency is quite similar and is

$$E_{FDH}^t(u^t, x^t | M^t, N^t) = \lambda_{FDH}^{t*}. \tag{10}$$

This measure is less than or equal to one. Again, a value of one deems a particular firm to be completely efficient.

2.3 Comparison

It has been well documented that the three forms of DEA are quite similar, but they collectively differ from the FDH technique. The main difference between the two approaches is how they define an efficient production set. The linear programming technique used in DEA attempts to measure the distance from the frontier of a convex envelope of the data. While the first two assumptions deal with dominance, the third may deem an undominated production plan inefficient because it does not lie on the convex envelope. In contrast, the FDH is more concerned with dominance than with distance. It only deems production plans which are dominated by other production plans to be inefficient.[7] Thus, the number of firms deemed efficient by the FDH is greater than or equal to those by each of the DEA models.

The difference between these formulations is also important from a managerial point of view. Some authors (see De Borger and Kerstens (1996), and Vanden Eeckaut, Tulkens and Jamar (1993)) argue that DEA calls inefficient too many observations because of the convexity assumptions. They argue that the assumptions, which often

---

[7]For a graphical representation see Fare, Grosskopf and Lovell (1994).

have no *a priori* support, fail to recognize local nonconvexities. Further, they attempt to discredit DEA models by stating that their comparison technique is flawed. Specifically due to the convexity assumptions, an inefficient observation is often compared to an unobservable and fictitious linear combination of efficient observations. Hence, it is illogical to claim that a firm can be dominated by another which does not exist. In contrast, the FDH reference technology is not vulnerable to this critique. It relates each inefficient observation to a single dominating observation.

On the other hand, others feel that the FDH allows for too many efficient observations. In the FDH framework, an observation with epsilon amount less of a particular input and a substantial amount less of output than an efficient firm may be deemed efficient, whereas that firm would be considered to be highly inefficient by DEA.

2.4   Statistical Inference

Discussion as to whether one methodology dominates another is not the only argument surrounding deterministic frontiers. The fact that the frontiers are deterministic is considered to be its main flaw by its skeptics. Many, including some econometricians, disapprove of the DEA and FDH type measures because they have been assumed to be non-stochastic and thus lacking statistical properties.

Recent advances in the literature model these so-called deterministic estimators in a statistical sense. Simar and Wilson (2001) use a bootstrapping method to assess the uncertainty about distance to the true production frontier from a relatively small number of points in the production set. In addition, they estimate confidence intervals for each observation. They do note that, in the case of the confidence intervals, samples containing many hundreds of thousands of observations would likely be infeasible due to the bootstrapping technique used. Although the usefulness of statistical inference

in deterministic frontiers to measure efficiency in panel data models appears to be promising, there needs to be significant work in this area.[8]

3   Stochastic Frontiers

Apart from some of the recent deterministic frontier research, the primary method of estimating a stochastic frontier is through econometrics. In addition to the stochastic aspect, there are several other features which distinguish the two methods from one another. As is discussed in Heshmati (1994), in the deterministic framework, the variations in firms performance are all attributed to inefficiency. This is problematic because often measurement errors, omitted variables and exogenous shocks are lumped into that measurement. Alternatively, in the econometric approach, by placing parameters in the error term, these and other effects can be distinguished from inefficiency. Another primary difference is that generally the econometric model specifies a specific functional form. The quality of the efficiency measurements are highly dependent on whether or not the functional form represents the true model. Gong and Sickles (1992) show that the stochastic model outperforms the DEA model, which needs no *a priori* specification of functional form, when the econometric model employed is close to the given underlying technology.[9]

Although little effort is being put forth into capturing a true technology, much re-

---

[8]Simar and Wilson (2000) and (2001) offer an in depth survey of this subject.

[9]Gong and Sickles (1992) use monte carlo experiments to compare DEA with several econometric techniques. They summarize the performance of the estimators by using the correlation and rank correlation coefficient between the true and estimated inefficiencies. These measure the degree to which deterministic and stochastic efficiency estimates are in accordance with the true data generating process (DGP). In the stochastic frontiers, the choice of functional form in characterizing an underlying technology is shown to be important in deriving unbiased information about the firm-specific technical efficiency. Further, the experimental evidence suggests that the tracking ability of either the deterministic or stochastic techniques deteroriates rapidly as the true technology becomes more complex.

search is devoted to finding the correct form of the error component. This section of the proposal will examine both time variant and time invariant measures of technical efficiency. Each topic will include estimation by random effects (RE), fixed effects (FE) and MLE. Following that, there will be a discussion on how heteroskedasticity affects each of these cases, and finally a discussion on the recent developments.

3.1 Time Invariant

The original time invariant efficiency model is written as

$$y_{it} = \alpha + \sum_{k} \beta_k x_{kit} + \varepsilon_{it}, \tag{11}$$

where

$$\varepsilon_{it} = v_{it} - u_i \tag{12}$$

$$i = 1, 2, ..., N$$

$$t = 1, 2, ..., T$$

$$k = 1, 2, ..., K,$$

where i indexes the firms, t indexes the time periods and k indexes the inputs. The endogenous variable y_{it} is output and the exogenous variables x_{kit} are K different inputs. The v_{it} random variables are assumed to be independent, uncorrelated with the regressors and often assumed to be normally distributed. Technical inefficiency is represented by u_i. The properties of u_i, other than it being non-negative, are determined by the specific model.

3.1.1 Random Effects

In the RE model, the u_i are assumed to be randomly distributed with a constant mean and variance. They are also assumed to be independent of both the random

11

errors and the regressors. Being that inefficiency can only take non-negative values, the distribution of u_i is often assumed to be half-normal, truncated normal, gamma or exponential.

Since the mean of u_i is often non-zero, estimation of technical efficiency is biased without the following modification for the RE model

$$
\begin{aligned}
\alpha^* &= \alpha - E(u_i) \\
&= \alpha - \mu
\end{aligned}
$$

$$
\begin{aligned}
u_i^* &= u_i - E(u_i) \\
&= u_i - \mu,
\end{aligned}
$$

where the u_i^* are iid with mean zero. The new model becomes

$$
y_{it} = \alpha^* + \sum_k \beta_k x_{kit} + v_{it} - u_i^*. \tag{13}
$$

Estimation can now be implemented using the standard two step generalized least squares (GLS) method. The values of u_i^* are estimated by the residuals using the feasible generalized least squares (FGLS) estimates of α^* and β_k^*, thus

$$
\widehat{u}_i^* = \widehat{\theta} \sum_t \widehat{v}_{it}, \tag{14}
$$

where

$$
\widehat{\theta} = \left(\frac{\widehat{\sigma}_u^2}{\widehat{\sigma}_v^2 + T\widehat{\sigma}_u^2} \right), \tag{15}
$$

12

and

$$\widehat{v}_{it} = y_{it} - \widehat{\alpha}^* - \sum_k \widehat{\beta}_k^* x_{kit}. \tag{16}$$

The estimates of u_i are obtained by

$$\widehat{u}_i = \max_i \widehat{u}_i^* - \widehat{u}_i^*. \tag{17}$$

This formulation ensures that at least one observation is considered to be efficient. These estimates are consistent as N and T tend towards infinity.

  Technical efficiency is then measured in one of two ways: if the production function is defined in levels, then Battese and Coelli (1988) state that technical efficiency for the *ith* firm should be measured (for a single input) by

$$TE_i = (\overline{x}_i \widehat{\beta} - \widehat{u}_i)(\overline{x}_i \widehat{\beta})^{-1}, \tag{18}$$

alternatively, if the production function is defined in terms of logarithms then the suggested measure is

$$TE_i = \exp(-\widehat{u}_i), \tag{19}$$

the values of which obviously range between 0 and 1. It should be noted that GLS is most appropriate when N is large because consistent estimation of σ_u^2 requires N to go to infinity. Schmidt and Sickles (1984) show that when N is small, GLS is useless unless σ_u^2 is known *a priori*. They also illustrate that when both N and T are large, GLS is feasible, but less efficient than the within estimator.

3.1.2 Fixed Effects

The technique used with the within estimator requires that u_i be fixed. This sets the basis for the FE model. The model is similar to the RE model except now the u_i term is subtracted from the overall intercept. The model becomes

$$y_{it} = \alpha_i + \sum_k \beta_k x_{kit} + v_{it}, \tag{20}$$

where

$$\alpha_i = \alpha - u_i. \tag{21}$$

In this model, the u_i are again assumed to be non-negative, but no independence or distributional assumptions are made. The within estimator is obtained by applying the within transformation, in which each variable is expressed in terms of deviations from its mean

$$y_{it} - \overline{y}_i = \sum_k \beta_k (x_{kit} - \overline{x}_{ki}) + v_{it}. \tag{22}$$

The N intercepts $(\alpha - u_i)$ are recovered from the residuals. Then the u_i are estimated by

$$\widehat{u}_i = \widehat{\alpha}_i - \max_i \widehat{\alpha}_i. \tag{23}$$

Obviously, the $\widehat{u}_i$ are non-negative and firm specific technical efficiency can again given by (18) or (19).

One of the drawbacks of the FE model is that although the estimates of the β_n's are consistent as either N or T go to infinity, the estimates of the α_i's are only consistent as T tends towards infinity. This is often not the case in economic panel data sets.

In addition, even in the case that T does tend towards infinity; it makes little sense to assume that technical efficiency will remain time invariant. Further, not only does the FE model capture the variation across producers, but it also captures the effects of all phenomenon (regulatory environment) that vary across producers, but are time invariant for each producer.

There are also advantages to choosing the FE model. Kumbhakar and Lovell (2000) show that besides its simplicity benefit, consistency does not depend on the distribution or independence of the variables. This could be potentially important since it seems quite possible that if a firm knows its level of inefficiency, its input values may be affected by that knowledge.[10]

3.1.3 Maximum Likelihood Estimation

The previous methods allow one to avoid strong distributional or independence assumptions. However, when these assumptions are reasonable within a particular panel, MLE is plausible.[11] Although several distributional assumptions can be made for the error terms, the following will be used throughout the rest of this paper for MLE cases

$$v_{it} \quad \sim \quad iid \ N(0, \sigma_v^2)$$

$$u_i \quad \sim \quad iid \ N^+(\mu, \sigma_u^2),$$

---

[10]A test may determine which model is appropriate for a particular panel data set. The Hausman-Taylor test allows one to test the uncorrelated hypothesis. This test by Hausman and Taylor (1981), based on the test by Hausman (1978), tests the significance between the within estimator and the GLS estimator.

[11]Refer to Kumbhakar (1987) or Battese and Coelli (1988) for a more in depth analysis.

where v_{it} and u_i are independent of one another and the regressors.[12] The log likelihood function for the normal-truncated normal case is

$$
\ln L = c - \frac{N(T-1)}{2} \ln \sigma_v^2 - \frac{N}{2} \left(\sigma_v^2 + T\sigma_u^2 \right) - N \ln \left[1 - \Phi \left(-\frac{\widetilde{\mu}}{\sigma_u} \right) \right] \qquad (24)
$$
$$
+ \sum_i \ln \left[1 - \Phi \left(-\frac{\widetilde{\mu}}{\sigma_*} \right) \right] - \frac{\sum \varepsilon_i' \varepsilon_i}{2\sigma_v^2} - \frac{N}{2} \left(\frac{\mu}{\sigma_u} \right)^2 + \frac{1}{2} \sum_i \left(-\frac{\widetilde{\mu}_i}{\sigma_*} \right)^2,
$$

where

$$
\widetilde{\mu}_i = \frac{\mu\sigma_v^2 - T\overline{\varepsilon}_i \sigma_u^2}{\sigma_v^2 + T\sigma_u^2}
$$
$$
\sigma_*^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma_v^2 + T\sigma_u^2}
$$
$$
\overline{\varepsilon}_i = (T)^{-1} \sum_t \varepsilon_{it},
$$

where $\widetilde{\mu}_i$ is a weighted mean for each i, σ_*^2 is a weighted variance term and $\overline{\varepsilon}_i$ is the average of ε_{it}. The conditional distribution of u given ε is

$$
f(u|\varepsilon) = \frac{f(u,\varepsilon)}{f(\varepsilon)}
$$
$$
= \frac{\exp\left(-\frac{(u-\widetilde{\mu})^2}{2\sigma_*^2} \right)}{(2\pi)^{\frac{1}{2}} \sigma_* \left[1 - \Phi \left(-\frac{\widetilde{\mu}}{\sigma_*} \right) \right]},
$$

where Φ denotes the distribution function of the standard normal random variable. After these have been derived, then either the mean or mode of the distribution can be used to help define producer specific estimates of time invariant technical efficiency. The mean and mode are listed as

$$
E(u_i|\varepsilon_i) = \widetilde{\mu}_i + \sigma_* \left(\frac{\phi\left(-\frac{\widetilde{\mu}}{\sigma_*} \right)}{1 - \Phi\left(-\frac{\widetilde{\mu}}{\sigma_*} \right)} \right), \qquad (25)
$$

---

[12] $N^+(\mu, \sigma_u^2)$ defines a truncated normal distribution with mean μ, and variance σ_u^2.

where ϕ represents the density function for the standard normal random variable, and

$$M(u_i|\varepsilon_i) = \{ \begin{matrix} \widetilde{\mu}_i \text{ if } \widetilde{\mu}_i \geq 0 \\ 0 \text{ otherwise} \end{matrix} \tag{26}$$

respectively. These values are again substituted into (18) or (19) to determine producer specific technical efficiency.[13]

3.2 Time Variant

Schmidt (1988) shows that time invariant technical efficiency is an attractive assumption, but on the other hand is a restrictive one. It seems quite implausible to assume that technical efficiency would remain constant over a prolonged period of time when the environment is competitive. When the panels are short, it may make sense to assume time invariant technical efficiency. Panels now have increased time length and when a "satisfactory" size panel is available, there seems to be little reason to assume u_i to be time invariant.

The time variant model is written like (11), with a similar error component, except that the u term is now indexed with t as well as i. The equations are again combined as in (20). The time variant model is

$$y_{it} = \alpha_{it} + \sum_k \beta_k x_{kit} + v_{it}, \tag{27}$$

where

$$\alpha_{it} = \alpha - u_{it}. \tag{28}$$

---

[13]It should be noted that MLE is generally more efficient than the other methods of estimation due to its exploitation of the distributional assumptions. See Kumbhakar and Lovell (2000) for more details on MLE and its comparison to the other two methods.

The problem with this specification is that with an $N \times T$ panel, it is impossible to estimate all of the $N \cdot T$ intercepts, the K slopes and σ_v^2. To avoid this problem, Cornwell, Schmidt and Sickles (1990) replace α_{it} with a flexible parameterized function of time with parameters that vary over time. The quadratic form of this is

$$\alpha_{it} = \theta_{i1} + \theta_{i2}t + \theta_{i3}t^2. \tag{29}$$

As a result, only $N \cdot 3$ intercepts need to be estimated with this setup. Additionally, the ratio of parameters to be estimated to the number of observations is $\frac{(3N+K+1)}{N}T$.

3.2.1 Random Effects

The RE model is estimated in almost exactly the same manner as the time invariant case. The GLS estimator is used and consistency hinges on the uncorrelatedness of u, v and the regressors. For a large T, it has the same properties as the time invariant model and is less efficient than the FE method.

3.2.2 Fixed Effects

The FE model has two methods for obtaining technical efficiency depending on the size of $\frac{N}{T}$. If the ratio is relatively large, then the u_{it}'s are deleted from the above equation. The slopes are estimated from the residuals, and the residuals are regressed on a constant, t and t^2 to obtain the estimates of θ_{i1}, θ_{i2} and θ_{i3}. This procedure will produce a value for $\widehat{\alpha}_{it}$ being

$$\widehat{\alpha}_{it} = \widehat{\theta}_{i1} + \widehat{\theta}_{i2}t + \widehat{\theta}_{i3}t^2. \tag{30}$$

Alternatively, if $\frac{N}{T}$ is relatively small, then the u_{it}'s are included in the model. In this case, the parameters of (29) are estimated as the coefficients of dummies. This will

18

give a similar estimated form of the intercepts. The estimated intercepts determine $\widehat{u}_{it}$, which is equal to

$$\widehat{u}_{it} = \max_i \widehat{\alpha}_{it} - \widehat{\alpha}_{it}. \tag{31}$$

Finally, technical efficiency for a specific firm in period t, in the logarithmic case, is defined to be

$$TE_{it} = \exp(-\widehat{u}_{it}), \tag{32}$$

and analogously in levels is defined to be (for a single regressor)

$$TE_{it} = (\overline{x}_i\widehat{\beta} - \widehat{u}_{it})(\overline{x}_i\widehat{\beta})^{-1}. \tag{33}$$

3.2.3 Alternative Formulations

There are alternative formulations for modeling the time varying u_{it}. Lee and Schmidt (1993) specify u_{it} as

$$u_{it} = \alpha(t)u_i, \tag{34}$$

where $\alpha(t)$ is a function of time dummy variables. It is obvious that the initial model is a special case of (34). Kumbhakar (1990) specifies a form of $\alpha(t)$ to be

$$\alpha(t) = (1 + \exp(\gamma t + \delta t^2))^{-1}, \tag{35}$$

where

$$0 \leq \alpha(t) \leq 1$$

and $\alpha(t)$ can be monotonically increasing or decreasing, concave or convex depending on the signs and magnitudes of the parameters. Kumbhakar (1990) and Battese and Coelli (1992) specify a form of $\alpha(t)$ to be

$$\alpha(t) = \exp(-\gamma(t - T)), \tag{36}$$

where $\alpha(t)$ is non-negative and decreasing at an increasing rate if $\gamma > 0$, increasing at an increasing rate if $\gamma < 0$ and constant if $\gamma = 0$.

3.2.4   Maximum Likelihood Estimation

Although some of these formulations are estimated with RE and FE methods, the majority are estimated though ML procedures. Again, if the independence and distribution assumptions are reasonable, the ML methods can be used. Note that the Kumbhakar (1990) and Battese and Coelli (1992) method will be discussed since it deals with the normal-truncated normal case. Further note that the steps are similar to those of the time invariant case. Log likelihood functions are derived, the conditional distributions are determined and the mean and mode for u_i given $\bar{\varepsilon}_i$ are

$$E(u_i|\varepsilon_i) = \mu_{*i} + \sigma_{**} \left(\frac{\phi\left(-\frac{\mu_{*i}}{\sigma_{**}}\right)}{1 - \Phi\left(-\frac{\mu_{*i}}{\sigma_{**}}\right)} \right) \tag{37}$$

and

$$M(u_i|\varepsilon_i) = \begin{cases} u_i^* \text{ if } \sum_t \alpha(t)\varepsilon_{it} \geq 0 \\ 0 \text{ otherwise} \end{cases} \tag{38}$$

where

$$\mu_{*i} = \frac{\mu\sigma_v^2 - \alpha'\bar{\varepsilon}_i\sigma_u^2}{\sigma_v^2 + \alpha'\alpha\sigma_u^2}$$

20

$$\sigma_{**}^2 = \frac{\sigma_u^2 \sigma_v^2}{\sigma_v^2 + \alpha'\alpha\sigma_u^2}$$

$$\alpha' = (\alpha(1), \alpha(2), ..., \alpha(T)),$$

where the variables μ_{*i} and σ_{**}^2 have similar interpretations as before. Note that the time invariant case is a special case of this setup when $\gamma = 0$. Further note that any of the previously defined α's can be used. Firm specific technical efficiencies are derived by placing (37) and (38) into (18) or (19), as previously discussed.[14]

3.3   Heteroskedasticity

The results of these models rely heavily on somewhat strict assumptions. As in the cross-sectional framework, these models assume that the variance of the error components are homoskedastic. In the classical linear regression model (CLRM), the Gauss-Markov assumptions require that the variance of the error term be constant. Many times the errors are not homoskedastic and the variances change drastically between observations. In the CLRM, heteroskedasticity causes estimates to be inefficient, although unbiased and consistent. Heteroskedasticity is possibly a larger problem in the error component model. It can appear in either component or in both, therefore affecting inferences dealing with technical efficiency. This section will analyze these cases in both the time invariant and time varying framework. When and if

---

[14]The following method of moments estimator can be found in Kumbhakar and Lovell (2000). The MM model is $\ln y_{it} = \alpha - \delta_t\sqrt{\frac{2}{\pi}}\sigma_n + \sum_k \ln x_{kit}'\beta_k + v_{it} - (u_{it} - E(u_{it}))$, this further decomposes to $\ln y_{it} = \delta_t^* + \sum_k \ln x_{kit}'\beta_k + v_{it} - u_{it}^*$. Where $u_{it} = u_i\delta_t$ and $E(u_{it}) = \delta_t\sqrt{\frac{2}{\pi}}\sigma_n$. To estimate the model, ordinary least squares estimation is first performed on the decomposed model with time dummies added. The coefficients of the time dummies are the δ_t^*'s. Then the residuals are used to compute the third moment for each t which are $m_{3t} = \delta_t^3 E(u_i - E(u_i))^3$. From this, δ_t is defined to be $\delta_t = \sigma_n^{-1}\left(\frac{m_{3t}}{\sqrt{\frac{2}{\pi}}}(1-\frac{4}{\pi})^{-1}\right)^{\frac{1}{3}}$. Normaliation of $\delta_1 = 1$ allows for the estimates of σ_n and δ_t to be obtained. These further allow for the estimation of α, δ_t and σ_v^2 which finally allow for the estimates of u_i, from which either the mean or mode of u_i conditional upon ε_i ultimately determines technical efficiency.

estimates of u_i and u_{it} are found, they can be used to estimate firm specific technical efficiency by using (18), (19) or (32) ,(33) as outlined before.

3.3.1 Time Invariant

Recall the previously discussed time invariant model (27). If heteroskedasticity is introduced into v_{it}, then the specification of the error component becomes

$$v_{it} \sim iid \ N(0, \sigma_{v_i}^2).$$

Independence is again assumed between the error components and between the error components and the regressors. Estimation is performed by GLS if $\sigma_{v_i}^2$ and σ_u^2 are known. If not, then they must be estimated. The estimation of σ_u^2 is the same as in homoskedastic case and $\sigma_{v_i}^2$ is estimated by

$$\widehat{\sigma}_{v_i}^2 = (T)^{-1} \sum_t \widehat{\varepsilon}_{it}^2,$$

where $\widehat{\varepsilon}_{it}$ are the residuals from the within transformed model. The estimates of u_i are

$$\widehat{u}_i = \max_i \overline{\varepsilon}_i - \overline{\varepsilon}_i, \tag{39}$$

where

$$\overline{\varepsilon}_i = (T)^{-1} \sum_t (y_{it} - \widehat{\alpha} - \sum_k \widehat{\beta}_k x_{kit}),$$

where $\widehat{\alpha}$ and $\widehat{\beta}_k$ are the GLS estimates.[15]

The FE approach is slightly different. The only assumptions are that

$$v_{it} \sim iid \ N(0, \sigma_{v_i}^2)$$

[15] A similar method for the cost function is presented in Kumbhakar (1997).

and that

$$u_i \geq 0.$$

The estimates for u_i are given by (23), the same as the homoskedastic case. It can be argued that failing to recognize heteroskedasticity in v_{it} may not be a serious problem in this situation.

This may not be the case in the MLE method. Although it is true that under certain conditions ML techniques are available, they are considered impractical unless $\frac{N}{T}$ is relatively small. If it is true that the ratio $\frac{N}{T}$ is relatively small, then the steps of estimation are similar to those in the time invariant section.[16]

Alternatively, when the u_i term is heteroskedastic, the majority of the methods are either impractical or impossible. The FE method is obviously impossible since the u_i term cannot be fixed and heteroskedastic. In the RE model, the ordinary least squares (OLS) procedure cannot be performed since the unconditional expectation of u_i (the intercept term $\alpha - E(u_i)$) is firm specific, the form of which depends on the distribution of u_i. This causes the OLS estimates to be biased and inconsistent. The MLE method is again impractical unless the ratio $\frac{N}{T}$ is relatively small.[17]

Similar difficulties occur for both the RE and FE methods when both v_{it} and u_i are heteroskedastic. Again, their estimators are impractical. MLE is plausible under certain conditions but is dominated by a methods of moments (MM) approach.

---

[16]Additional information including a method of moments estimator similar to the one discussed in a previous footnote can be found in Kumbhakar and Lovell (2000).

[17]The only alternative is the method of moments estimator described in Kumbhakar and Lovell (2000).

3.3.2 Time Variant

Heteroskedasticity becomes an even bigger problem in the time varying case. Recall that the model of Kumbhakar (1990) and Battese and Coelli (1992), it is still of the form (27) and α_{it} is again measured by (34). Further, v_{it} and u_{it} are assumed to be random. Many difficulties arise when attempting to estimate the model under a heteroskedastic v_{it}. Under the usual assumptions of the error components, the ML approach may be impractical if either N or T are large. This is because there are N variance parameters and $T-1$ efficiency parameters to be estimated. Although possible, the ML approach is dominated by a MM estimator. As in the previous section, when u_i is heteroskedastic or when both v_{it} and u_i are heteroskedastic, the ML estimator is impractical and a MM estimator dominates.[18]

3.4   Recent Developments

Currently, many extensions are being posed to improve the estimation and measurement of technical efficiency. Some of these are being accepted, others are not. This section briefly reviews three of these extensions.

3.4.1   Firm Specific and Time Specific Effects

The imposition of time varying technical efficiency is not sufficient for some authors. Kumbhakar (1991) devises a way to separate firm specific and time specific effects from technical efficiency. Kumbhakar argues that without these effects, the estimates of u_{it} are biased. The model again is

---

[18]The steps of which are only slightly different from the time invariant case found in Kumbhakar and Lovell (2000).

$$y_{it} = \alpha + \sum_k \beta_k x_{kit} + \varepsilon_{it}, \tag{40}$$

but now where

$$\varepsilon_{it} = \mu_i + \lambda_t + v_{it} - u_{it}, \tag{41}$$

where μ_i is the firm specific effect, λ_t is the time specific effect and the rest of the variables have their usual interpretations. MLE can be implemented with the usual distribution and independence assumptions.[19]

3.4.2 Bayesian

Another recent development in the literature is that of Bayesian estimation. The basic panel data stochastic frontier model (11) is the same. The FE approach yields the same estimates of u_i (23) and firm specific technical efficiency is again measured by (18) or (19). The only problem with this approach, as pointed out by Koop, Osiewalski and Steel (1997), is that the FE model favors low efficiency. In contrast, the Bayesian RE model has an advantage such that it can distinguish u_i from the intercept. In addition it is also possible to estimate absolute efficiency as opposed to relative efficiency.[20]

3.4.3 Long Run Inefficiency Levels

Another interesting advancement is that of the study of long run (LR) efficiency levels. Ahn, Good and Sickles (2000) take a dynamic frontier approach. Their approach not

---

[19]Those interested in a futher discussion, including estimation, should consult Heshimati and Kumbhakar (1994).

[20]Kim and Schmidt (2000) offer a more in-depth description and a comparison between Bayesian and Classical approaches.

only differs from the previous models because of the dynamic portion, but also because past models only analyzed short run (SR) dynamics. The model for the LR at first glance appears to be the same as (27), except now the u_{it} term is defined as

$$u_{it} = (1 - \rho_i)u_{i(t-1)} + \eta_{it}, \tag{42}$$

where

$$E(\eta_{it}) = \kappa_i > 0.$$

Each firms inefficiency follows an autoregressive of order one (AR(1)) process, where ρ_i $(0 \leq \rho_i \leq 1)$ measures firm i's ability to adjust its last periods inefficiency level. η_{it} is a non-negative random noise and κ_i is a positive constant for each i.[21]

4 Conclusion

This chapter surveys a wide variety of methods to measure technical efficiency from panel data. First were the deterministic approaches of DEA and the FDH. The discussion regarding the difference between the two approaches causes this author to conclude that although there are some strict assumptions in DEA, it appears to have a more intuitive result. In addition, if a continuum of firms existed it seems plausible that the FDH and DEA frontiers would be identical. Following that was a brief look at the statistical inference methods in deterministic frontiers. This appears to be the next major area of research within the deterministic frontiers.

Next were the stochastic approaches estimated with econometrics. These focused on both time invariant and time variant technical efficiency. Models were estimated by RE, FE and MLE. In addition, heteroskedasticity was introduced into the error

---

[21]Estimation and an empirical example can be found in Anh, Good and Sickles (2000).

components and analyzed by each method. The FE models appear to be romising due to their generality, but RE models often perform best with "typical" economic panel data.

Finally, extensions from these models were presented. Perhaps more interesting than those, would be an attempt to merge the benefits of deterministic and stochastic frontiers. Applying nonparametric kernel methods to the stochastic frontier literature could give the flexibility enjoyed by DEA and FDH as well as the stochastic element beloved by econometricians. Although no conclusions can be drawn from this thought, an answer to whether one method dominates the other may be obtained through such an exercise.

CHAPTER 2
A NONPARAMETRIC RANDOM EFFECTS
ESTIMATOR[1]

1 Introduction

A nonparametric stochastic panel data estimator is desireable, but which method should be undertook? The random effects approach treats u_i as a group specific disturbance.. This framework is most appropriate when the cross-sectional units are believed to be sampled from a large population. The fixed effects approach treats the individual effect as a group specific constant term within the regression model. This approach is most appropriate when the cross-sectional units are the complete set of the population, meaning the researcher can be confident that the differences between the cross-sectional units can be viewed as parametric shifts of the regression function. Maximum likelihood estimation treats the u_i as random disturbances that follow a particular distribution. This approach is most appropriate when the distribution of the individual effect is known. Since most economic data is a sample taken from a larger population, in this chapter we consider the estimation of random effects models where u_i is random.[2]

In the case of linear regressions, a particular concern has been with the linearity of the functional form connecting the variables of the model. Often the true technology is unknown and linear regressions are performed without economic reasoning due to their straightforward estimation procedures and well-known properties. This concern initially spawned an interest in transformations of the endogenous and exogenous

---

[1] This chapter is taken from Henderson and Ullah (2003).

[2] For further discussion and tests to detmine which method is appropiate as well as estimation by the other techniques, see Greene (2002).

variables, leading to the use of flexible specifications, such as the translog functional form. Although approaches such as these have served econometrics well, there has always been some worry that the functional form might be more complex. Thus, it is worthwhile considering nonparametric estimation if the functional form is unknown. The basic idea behind nonparametric estimation is to approximate the technology arbitrarily close. Unfortunately, the literature, up to this point does not possess a nonparametric kernel estimator for the one-way error component random effects model. Thus, in order to develop a nonparametric method to estimate technical efficiency, a nonparametric random effects estimator must first be developed. This chapter presents such an estimator, one in which no functional form is associated with any of the regressors.[3]

This chapter is organized as follows: Section 2 gives the model, notation, proposes a new estimator and derives the theoretical estimates. Section 3 provides the Monte Carlo setup and summarizes the results of the experiments. Finally, the fourth section concludes.

2 The Model

Let us consider a nonparametric one-way error component model as

$$y_{it} = m(x_{it}) + \varepsilon_{it}, \tag{1}$$

where $i = 1, 2, ..., N$, $t = 1, 2, ..., T$, y_{it} is the endogenous variable, x_{it} is a vector of k exogenous variables and $m(\cdot)$ is an unknown smooth function. Further, ε_{it} follows the one-way error component specification

---

[3]For an example of semiparametric methods, (making use of the Robinson (1988) technique and examining the parametric parameters) see Berg, Li and Ullah (2000).

$$\varepsilon_{it} = u_i + v_{it}, \tag{2}$$

where u_i is $i.i.d.$ $(0, \sigma_u^2)$, v_{it} is $i.i.d.$ $(0, \sigma_v^2)$ and u_i and v_{it} are uncorrelated for all q and ls, where q is different from l; $q, l \in i$ and $s \in t$.

Let $\varepsilon_i = [\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{iT}]'$ be a $T \times 1$ vector. Then $V \equiv E(\varepsilon_i \varepsilon_i')$, takes the form

$$V = \sigma_v^2 \mathbf{I}_T + \sigma_u^2 \mathbf{i}_T \mathbf{i}_T', \tag{3}$$

where $\mathbf{I}$ is an identity matrix of dimension T and $\mathbf{i}$ is a $T \times 1$ column vector of ones. Since the observations are independent over q and l, the covariance matrix for the full $NT \times 1$ disturbance vector ε, $\Omega = E(\varepsilon \varepsilon')$ is

$$\Omega = V \otimes \mathbf{I}_N. \tag{4}$$

We are interested in estimating the unknown function $m(x)$ at a point x and the slope of $m(x)$, $\beta(x) = \nabla m(x)$, where ∇ is the gradient vector of $m(x)$. The parameter $\beta(x)$ is interpreted as a varying coefficient. We consider the usual panel data situation of large N and small T.

Nonparametric kernel estimation of $m(x)$ and $\beta(x)$ can be obtained by using local linear least squares (LLLS) estimation. This is obtained by minimizing the local least squares or weighted least squares of errors

$$\sum_i \sum_t (y_{it} - X_{it}\delta(x))^2 K\left(\frac{x_{it} - x}{h}\right) = (y - X\delta(x))K(x)(y - X\delta(x)) \tag{5}$$

with respect to $m(x)$ and $\beta(x)$, where y is a $NT \times 1$ vector, X is a $NT \times (k+1)$ matrix generated by $X_{it} = (1 \quad (x_{it} - x))$, $\delta(x) = (m(x), \beta(x))'$ is a $(k+1) \times 1$ vector, $K(x)$ is an $NT \times NT$ diagonal matrix of kernel (weight) functions $K(\frac{x_{it} - x}{h})$ and h is the

bandwidth (smoothing) parameter. Generally kernel functions can be any probability function having a finite second moment (here we use the standard normal kernel). The estimator so obtained is

$$\widehat{\delta}(x) = (X'K(x)X)^{-1}X'K(x)y \tag{6}$$

The estimator of $m(x)$ is then given by $\widehat{m}(x) = (1 \quad 0)\widehat{\delta}(x)$, whereas $\widehat{\beta}(x)$ can be extracted from $\widehat{\delta}(x)$ as $\widehat{\beta}(x) = (0 \quad 1)\widehat{\delta}(x)$. The estimator in (6) is called LLLS estimator. Asymptotic normality for the cross-sectional case is proven by Li and Woolridge (2000) and Kniesner and Li (2002) derive a proof for the case of panel data.[4]

The LLLS estimator in (6) however ignores the information contained in the disturbance vector covariance matrix Ω. In view of this we introduce a new estimator, local linear generalized least squares (LLGLS) estimator, by minimizing

$$(y - X\delta(x))'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}(y - X\delta(x)) \tag{7}$$

with respect to $\delta(x)$. This gives the estimator $d(x)$

$$d(x) = (X'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}X)^{-1}X'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}y. \tag{8}$$

The objective function in (7) amounts to doing GLS (linear) fits to the points local to x.[5] The LLGLS estimator in (8) however depends upon the unknown parameters

---

[4]For more information on the choices of K and h see Fan and Gijbels (1992) and Pagan and Ullah (1999).

[5]We also consider an alternative nonparametric estimator which takes the form: $\widehat{\delta}(x)_{ANPFGLS} = (X'\widehat{\Omega}^{-\frac{1}{2}}K(X)\widehat{\Omega}^{-\frac{1}{2}}X)^{-1}(X'\widehat{\Omega}^{-\frac{1}{2}}K(X)\widehat{\Omega}^{-\frac{1}{2}}Y)$. Although it looks similar, there is an inherent flaw in the way the Alternative Nonparametric Feasible Generalized Least Squares (ANPFGLS) estimator transforms the data. By simply looking at the matrix structure it becomes evident. It shows that in the ANPFGLS setup the variables are first adjusted for heteroskedasticity (i.e. $X'\widehat{\Omega}^{-\frac{1}{2}} = X^*$), and then run through the kernel (i.e. $X^{*'}K(X)$). The basic idea behind the nonparametric estimator is to provide a separate estimate for each value of x, the formulation of the ANPFGLS

σ_u^2 and σ_v^2. An estimator of σ_v^2 is obtained by using the within estimator. The exact form of this $\widehat{\sigma}_v^2$ is

$$\widehat{\sigma}_v^2 = \frac{1}{NT} \sum_i \sum_t \widehat{(v_{it} - \overline{v}_i)}^2 \tag{9}$$

where

$$\sum_i \sum_t \widehat{(v_{it} - \overline{v}_i)}^2 = \sum_i \sum_t [(y_{it} - \overline{y}_i) - (x_{it} - \overline{x}_i)\beta^*(x_{it})]^2 \tag{10}$$

in which $\overline{y}_i = \frac{1}{T} \sum_t y_{it}$, $\overline{x}_i = \frac{1}{T} \sum_t x_{it}$ and

$$\beta^*(x) = [\sum_i \sum_t (x_{it} - \overline{x}_i)'(x_{it} - \overline{x}_i)K_{it}]^{-1} \sum_i \sum_t (x_{it} - \overline{x}_i)'(y_{it} - \overline{y}_i)K_{it}, \tag{11}$$

where $K_{it} = K\left(\frac{x_{it} - x}{h}\right)$.

One estimate of σ_u^2 is obtained as a combination of $\widehat{\sigma}_v^2$ and $\widehat{\sigma}_\varepsilon^2$ being

$$\widehat{\sigma}_u^2 = \widehat{\sigma}_\varepsilon^2 + \frac{1}{T}\widehat{\sigma}_v^2, \tag{12}$$

in which

$$\widehat{\sigma}_\varepsilon^2 = \frac{1}{N - k} \sum_i \widehat{\overline{\varepsilon}}_i^2, \tag{13}$$

where

$$\sum_i \widehat{\overline{\varepsilon}}_i^2 = \frac{1}{N} \sum_i \left[(\overline{y}_i - \overline{y}) - (\overline{x}_i - \overline{x})\widetilde{\beta}(x_i)\right]^2, \tag{14}$$

where

gives a separate estimate for each value of x^*. Therefore it is this process of the variables being adjusted for heteroskedasticy before they are smoothed which causes the mispecification, whereas the NPFGLS estimators first transform the data correctly (i.e. $X'K(X)$) and then proceed to adjust it for heteroskedasticity.

$$\widetilde{\beta}(x) = \frac{\sum_i \left(\overline{y}_i - \overline{y}\right)\left(\overline{x}_i - \overline{x}\right) K\left(\frac{x_i - x}{h}\right)}{\sum_i \left(\overline{x}_i - \overline{x}\right)^2 K\left(\frac{x_i - x}{h}\right)}. \tag{15}$$

where $\overline{y} = \frac{1}{N}\sum_i \overline{y}_i$ and $\overline{x} = \frac{1}{N}\sum_i \overline{x}_i$.

Alternative estimators of σ_u^2 and σ_v^2 can be obtained by noting that $V(\varepsilon_{it}) \equiv \sigma_\varepsilon^2 = \sigma_u^2 + \sigma_v^2$ and

$$cov(\varepsilon_{qt}, \varepsilon_{lt'}) = \sigma_u^2 \text{ for } q \neq l$$

and zero otherwise. Thus

$$\widehat{\sigma}_\varepsilon^2 = \frac{1}{NT}\sum_i\sum_t \widehat{\varepsilon}_{it}^2 \tag{16}$$

and

$$\widehat{\sigma}_u^2 = \frac{1}{N(T-1)}\sum_i\sum_t\sum_{t\neq t'} \widehat{\varepsilon}_{it}\widehat{\varepsilon}_{it'}' \tag{17}$$

where $\widehat{\varepsilon}_{it} = y_{it} - X_{it}\widehat{\delta}(x_{it})$ is the LLLS residual based on the first stage estimator of $\widehat{\delta}(x)$ in (6).[6]

Substituting the estimators of σ_u^2 and σ_v^2 from (9) and (12) or (16) and (17) into (8) gives a feasible linear generalized least squares (FLGLS) or nonparametric feasible generalized least squares (NPFGLS) estimator as

$$\widehat{\delta}(x)_{NPFGLS} = (X'\sqrt{K(x)}\widehat{\Omega}^{-1}\sqrt{K(x)}X)^{-1}X\sqrt{K(x)}\widehat{\Omega}^{-1}\sqrt{K(x)}y.$$

Consistency of $\widehat{\delta}(x)$ is straightforward under the standard regularity conditions, but asymptotic normality requires somewhat stronger assumptions as stated in Li and

---

[6]Note that estimation of the third variance term is straightforward. Also, it has been noted by Maddala and Mount (1973) and Taylor (1980) that more efficient esimation of the variance components does not necessarily lead to more effecient estimates of $m(x)$ and $\beta(x)$.

Woolridge (2000, p. 340). Asymptotic normality is established under the following theorem.

THEOREM:

Under the standard assumptions

$$D(NT)\left(\widehat{\delta}(x) - \delta(x) - \begin{pmatrix} h^2 \mu_k W \\ 0 \end{pmatrix}\right) \to N(0, \Sigma_x)$$

where $D(NT) = \begin{pmatrix} (NTh^k)^{\frac{1}{2}} & 0 \\ 0 & (NTh^{k+2})^{\frac{1}{2}} \end{pmatrix}$, $\mu_k = \frac{1}{2}c_k tr(m''(x))$, $W = tr(\Omega^{-1})$,

$\Sigma_x = \begin{pmatrix} d_k \sigma_\varepsilon^2 / f(x) W & 0 \\ 0 & v_k \sigma_\varepsilon^2 I_k / c_k^2 f(x) W \end{pmatrix}$, $c_k = \int K(\psi)\psi\psi' d\psi$, $d_k = \int K^2(\psi)d\psi$ and

$v_k = \int K^2(\psi)\psi\psi' d\psi$, where ψ is defined as $\frac{x_{it}-x}{h}$. The proof is given in the appendix. It loosely follows the method used by Li and Wooldridge (2000) for the cross-sectional case. The proof differs by incorporating information about the variance parameters. Also note that the proof is given for the most general case (LLGLS) and is easily modified to satisfy the above scenario.

3   Monte Carlo Results

Although asymptotic results give clues as to the performance of the estimators, most economic panel data is finite. This section uses Monte Carlo simulations to examine the finite sample performance of the proposed NPFGLS estimator. Following the methodology of Baltagi, Chang and Li (1992), the following data generating process is used:

$$y_{it} = \alpha + x_{it}\beta + x_{it}^2 \gamma + u_i + v_{it},$$

where x_{it} is generated by the method of Nerlove (1971).[7] The value of α is chosen to be 5, β is chosen to be 0.5 and γ takes the values of 0 (linear technology) and

---

[7]The x_{it} were generated as follows: $x_{it} = 0.1t + 0.5x_{it-1} + w_{it}$, where $x_{i0} = 10 + 5w_{i0}$ and $w_{it} \sim U[-\frac{1}{2}, \frac{1}{2}]$.

2 (quadratic technology). The distribution of u_i and v_{it} are generated separately as i.i.d. Normal. Total variance of $\sigma_v^2 + \sigma_u^2 = 20$ and $\rho = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ is varied to be 0.1, 0.4 and 0.8.

For comparison, we compute the following estimators of δ:

(I) Parametric (linear) Feasible GLS (FGLS) estimator

$$\widehat{\delta}_{FGLS} = (X'\widehat{\Omega}^{-1}X)^{-1}(X'\widehat{\Omega}^{-1}Y).$$

(II) NPFGLS estimator

$$\widehat{\delta}_{NPFGLS} = (X'\sqrt{K(x)}\widehat{\Omega}^{-1}\sqrt{K(x)}X)^{-1}X\sqrt{K(x)}\widehat{\Omega}^{-1}\sqrt{K(x)}y.$$

Reported are the estimated bias and mean squared error (MSE) for each estimator. These are computed via $Bias(\widehat{m}) = M^{-1} \sum_j (\widehat{m}_j - m^*)$, and $MSE(\widehat{m}) = M^{-1} \sum_j (\widehat{m}_j - m^*)^2$ where M is the number of replications, $\widehat{m}_j$ is the estimated value of m^* at the jth replication and $m^* = \alpha + \overline{x}\beta + \overline{x}^2\gamma$. Similarly for the varying coefficient parameter, $Bias(\widehat{\beta}) = M^{-1} \sum_j (\widehat{\beta}_j - \beta^*)$ and $MSE(\widehat{\beta}) = M^{-1} \sum_j (\widehat{\beta}_j - \beta^*)^2$, where $\widehat{\beta}_j$ is the estimated value of β^* at the jth replication and $\beta^* = \beta + 2\overline{x}\gamma$. $M = 1000$ is used in all simulations, T is varied to be 3 and 5, while N takes the values 10, 20 and 50. The simulation results are given in Tables 1 through 4. The smallest MSE for each case (for a given N, T, ρ and γ) is shown as a boldface number.[8]

Tables 1 and 2 report the result for $\gamma = 0$ (linear technology). In each case, the parametric estimators outperform the nonparametric estimators in MSE. This result

---

[8]In the tables only local estimates of $\widehat{\delta}$ are given. The specific form being

$$\widehat{\delta}(\overline{x})_{NPFGLS} = (X'K(\overline{x})^{\frac{1}{2}}\widehat{\Omega}^{-1}K(\overline{x})^{\frac{1}{2}}X)^{-1}XK(\overline{x})^{\frac{1}{2}}\widehat{\Omega}^{-1}K(\overline{x})^{\frac{1}{2}}y,$$

where $K(\overline{x}) = I_{NT} \times K\left(\frac{x_{it}-\overline{x}}{h}\right)$. These estimates are similar to the global ones, except that rather than being evaluated at each value of x and then averaged, they are only evaluated at the mean value of x. Being that these estimates are evaluated at the mean of x, they appear to perform better in Monte Carlo exercises that evaluate at the mean of x. It should be noted however that this method rarely outperforms the global measures in empirical data and is usually only used for computational ease. Further, replacing these with the global estimates does not affect the conclusions of the paper.

is expected since the true underlying technology is linear and because nonparametric estimators are known to have small sample bias. Contrary to the first two tables, the results of Tables 3 and 4 are in support of the NPFGLS estimators. The NPFGLS estimators outperform the parametric estimators drastically in both $Bias$ and MSE. This table shows the major drawback of the parametric type estimators and the strengths of the nonparametric type estimators. When the technology is complex, parametric estimators are often misspecified and the nonparametric estimators, with their complete flexibility, are better able to adapt.

4   Conclusion

This chapter examines the problem of improving the estimation of a one-way random effects error component model. A nonparametric estimator is proposed, its structure is defined, asymptotic properties proven and its finite sample results are generated through a Monte Carlo exercise. The Monte Carlo results of section 3 show that the parametric estimators provide slightly smaller MSE when the true underlying technology is linear and correctly specified. Although less efficient in the aforementioned exercises, the NPFGLS estimators provide acceptable estimation. On the other hand, when the technology becomes complex, the NPFGLS estimators perform best, with the MSE of the linear parametric estimator up to 70 times that of the NPFGLS estimator. Thus, it is suggested that the NPFGLS estimators be used in practice because the true underlying technology is usually unknown; and as Tables 3 and 4 demonstrate, the consequences of choosing the wrong estimator may be quite severe.

CHAPTER 3
NONPARAMETRIC KERNEL MEASUREMENT OF TECHNICAL EFFICIENCY

1 Introduction

Given the results in the previous chapter, it is now possible to derive a nonparametric method to estimate technical efficiency. The objective of this chapter is to provide nonparametric estimates of production frontiers and of technical efficiency. Further, the sampling properties of these estimators are analyzed. It is shown that the nonparametric estimates provide improved estimation compared to misspecified parametric technical efficiency measures.

This chapter is organized as follows: Section 2 gives the model and notation, proposes the nonparametric estimators and derives their theoretical estimates. The third section examines an alternative method for constructing nonparametric production frontiers. Section 4 provides the Monte Carlo setup and summarizes the results of the experiments. The fifth section provides an empirical example involving cross-country data used in Fare, Grosskopf, Norris and Zhang (1994). Finally, Section 6 concludes.

2 The Model

The nonparametric random effects production frontier model takes the form

$$y_{it} = g(x_{it}) + \varepsilon_{it}, \tag{1}$$

where $i = 1, 2, ..., N$, $t = 1, 2, ..., T$, $y_{it} \geq 0$ is output, $x_{it} \geq 0$ is a vector of k inputs and $g(\cdot)$ is an unknown smooth production function. The ε_{it} follow the production frontier error component specification

$$\varepsilon_{it} = v_{it} - u_i, \tag{2}$$

where

$$u_i \geq 0 \tag{3}$$

for all i. The u_i represent producer specific, time invariant inefficiency and are $i.i.d.$ (μ, σ_u^2), v_{it} are the random distribuances and are $i.i.d.$ $(0, \sigma_v^2)$ and u_i and v_{it} are uncorrelated for all q and ls, where q is different from l; $q, l \in i$ and $s \in t$. Since the mean of u_i is often non-zero, estimation of technical efficiency is biased without the following modification to the random effects model:

$$y_{it} = m(x_{it}) + v_{it} + u_i^*, \tag{4}$$

where

$$m(x_{it}) = g(x_{it}) - \mu \tag{5}$$

and

$$u_i^* = \mu - u_i, \tag{6}$$

where, obviously, the u_i^* are distributed i.i.d. with a zero mean.

As in all random effect (Generalized Least Squares) type estimators, knowledge of the error variance structure is necessary. Letting $\varepsilon_i = [\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{iT}]'$ and defining $V \equiv E(\varepsilon_i \varepsilon_i')$ gives

$$V = \sigma_v^2 \mathbf{I}_T + \sigma_u^2 \mathbf{i}_T \mathbf{i}_T', \tag{7}$$

38

where $\mathbf{I}$ is an identity matrix of dimension T and $\mathbf{i}$ is a $T \times 1$ column vector of ones. The independence assumption of observations over q and l gives the disturbance covariance matrix for the full NT observations as

$$\Omega = V \otimes \mathbf{I}_N. \tag{8}$$

Estimation of $m(x)$, $\beta(x)$ (the slope of $m(x)$) and u_i^* can now be performed via a two step procedure. First, estimation of $m(x)$ and $\beta(x)$ are obtained by the Nonparametric Feasible Generalized Least Squares (NPFGLS) estimator. Again, the estimator, in matrix form, is

$$\widehat{\delta}(x)_{NPFGLS} = (X'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}X)^{-1}X'\sqrt{K(x)}\Omega^{-1}\sqrt{K(x)}y, \tag{9}$$

where

$$X = (1 \quad x_{it}),$$

and $K(x) = I_{NT} \times K\left(\frac{x_{it}-x}{h}\right)$. $K\left(\frac{x_{it}-x}{h}\right)$ is the standard normal kernel function with optimal bandwidth h.[1] The estimate of $m(x)$ is extracted from $\widehat{\delta}(x)$ by $\widehat{m}(x) = (1 \quad 0)\widehat{\delta}(x)$, whereas $\beta(x)$ can be obtained from $\widehat{\delta}(x)$ as $\widehat{\beta}(x) = (0 \quad 1)\widehat{\delta}(x)$.

To obtain the estimates of u_i^*, the second step of the estimation is performed by maximizing the following objective function

$$S = \frac{1}{\sigma_v^2}\sum_i\sum_t (y_{it} - \widehat{m}(x_{it}) - u_i^*)^2 + \frac{1}{\sigma_u^2}\sum_i\sum_t u_i^{*2}, \tag{10}$$

with respect to u_i^*. The estimates of u_i^* come as

---

[1]See Pagan and Ullah (1999) for details on the properites of the standard normal kernel, as well as alternatives, and bandwidth selection.

$$\widehat{u}_i^* = \widehat{\theta} \sum_t \widehat{v}_{it}, \tag{11}$$

where

$$\widehat{\theta} = \left(\frac{\widehat{\sigma}_u^2}{\widehat{\sigma}_v^2 + T\widehat{\sigma}_u^2} \right), \tag{12}$$

and

$$\widehat{v}_{it} = y_{it} - \widehat{m}(x_{it}). \tag{13}$$

The estimate of σ_v^2 can again obtained using the within estimator. The exact form of $\widehat{\sigma}_v^2$ is

$$\widehat{\sigma}_v^2 = \frac{1}{NT - k - N} \sum_i \sum_t (\widehat{v_{it} - \overline{v}_i})^2, \tag{14}$$

where

$$\sum_i \sum_t (\widehat{v_{it} - \overline{v}_i})^2 = \sum_i \sum_t \left[(y_{it} - \overline{y}_i) - (x_{it} - \overline{x}_i)\beta^*(x) \right]^2. \tag{15}$$

and

$$\beta^*(x) = \left[\sum_i \sum_t (x_{it} - \overline{x}_i)(x_{it} - \overline{x}_i)' K_{it} \right]^{-1} \sum_i \sum_t (x_{it} - \overline{x}_i)'(y_{it} - \overline{y}_i)K_{it}, \tag{16}$$

where

$$K_{it} = K\left(\frac{x_{it} - x}{h} \right),$$

$$\overline{y}_i = \sum_t y_{it},$$

and

$$\overline{x}_i = \sum_t x_{it}.$$

A combination of $\widehat{\sigma}_v^2$ and $\widehat{\sigma}_\varepsilon^2$ can be used to estimate σ_u^2 as

$$\widehat{\sigma}_u^2 = \widehat{\sigma}_\varepsilon^2 + \frac{1}{T}\widehat{\sigma}_v^2, \tag{17}$$

where

$$\widehat{\sigma}_\varepsilon^2 = \frac{1}{N-k} \sum_i \widehat{\widetilde{\varepsilon}}_i^2, \tag{18}$$

where

$$\sum_i \widehat{\widetilde{\varepsilon}}_i^2 = \sum_i \left[(\overline{y}_i - \overline{y}) - (\overline{x}_i - \overline{x})\, \widetilde{\beta}(x) \right]^2, \tag{19}$$

where

$$\widetilde{\beta}(x) = \frac{\sum_i (\overline{y}_i - \overline{y})(\overline{x}_i - \overline{x}) K_i}{\sum_i (\overline{x}_i - \overline{x})^2 K_i} \tag{20}$$

where

$$K_i = K\left(\frac{\overline{x}_i - x}{h}\right),$$

$$\overline{y} = \frac{1}{N} \sum_i \overline{y}_i,$$

and

$$\overline{x} = \frac{1}{N} \sum_i \overline{x}_i.$$

Finally, the estimates of u_i^* can be used to estimate producer specific technical efficiency. First, the estimates of u_i are obtained by means of the normalization

$$\widehat{u}_i = \max_i \widehat{u}_i^* - \widehat{u}_i^*, \tag{21}$$

and estimates of technical efficiency can be defined as

$$\widehat{TE}_i = \exp(-\widehat{u}_i). \tag{22}$$

Again, this method, like others, assures that the values of technical efficiency are between zero and one. A value of one defines a firm as technically efficient and a value below one deems a firm as inefficient.

3   Monte Carlo Results

Since most economic panel data is considered finite, asymptotic results can only gives clues as to how the estimators will perform with real data. This section uses Monte Carlo simulations to examine the finite sample performance of the proposed estimators. Modifying the exercise used in Gong and Sickles (1992), the following data generating process is used:

$$y_{it} = \sum_k \gamma_k x_{kit}^{\alpha_k} + v_{it} - u_i, \tag{23}$$

where $\alpha_k > 0 \; \forall k$ and $\sum_k \gamma_k = 1$. The method used in Nerlove (1971) to generate x_{it} is modified to be consistent with a production technology.[2] For simplification, k is chosen to be 1 and α takes the values 1 (CRS), $\frac{1}{2}$ (NRS) and 2 (VRS). The distribution of u_i and v_{it} are generated as i.i.d half-normal, exponential, and gamma

---

[2]The x_{it} were generated as follows: $x_{it} = 0.1t + 0.5x_{it-1} + w_{it}$, where $x_{i0} = 10 + 5w_{i0}$ and $w_{it} \sim U[0,1]$.

(all non-negative distributions) and normal respectively.[3] Further, the total variance of $\sigma_v^2 + \sigma_u^2 = 20$ and $\rho = \sigma_u^2/(\sigma_u^2 + \sigma_v^2)$ is varied to be 0.1, 0.4 and 0.8.

For comparison, the following estimates of technical efficiency are obtained:

(I) Parametric[4]

$$TE_P = \exp(-\tilde{u}_i)$$

(II) NPFGLS

$$TE_{NP} = \exp(-\hat{u}_i)$$

(III) DEA- CRS

$$TE_{DEA-CRS} = \lambda_{CRS}^{t*}$$

(IV) DEA - NRS

$$TE_{DEA-NRS} = \lambda_{NRS}^{t*}$$

(V) DEA - VRS

$$TE_{DEA-VRS} = \lambda_{VRS}^{t*}$$

Reported are the correlations between the estimated efficiency values and the true efficiency values. These are calculated, for the parametric estimators (from which the other versions are obvious), as

$$CORR(TE, TE_P) = M^{-1} \sum_j \left(CORR_j \left(TE, TE_P \right) \right), \tag{24}$$

---

[3]Although these three distributions of u_i are used in most stochastic frontier studies, several authors have expressed that they may be less than desireable (see Carree (2002), Li (1996) and Ritter and Simar (1997)).

[4]The parametric estimates of δ, Ω, θ, v_{it} and u_i can be found in Kumbhakar and Lovell (1999).

where $TE = \exp(-u_i)$ are the true efficiency values, TE_P are the estimated values, $CORR(TE, TE_P)_j$ is the correlation between the two at the jth replication and M is the number of replications. $M = 1000$ is used in all simulations, T is varied to be 5 and 7, while N takes the values 10, 20 and 50. The simulation results are given in Tables 5 thru 13. The largest correlation coefficient for a given case (for a given N, T, ρ and α) is shown as a boldface number.[5]

The results for the estimators, when the the true underlying process is linear, are given in Tables 5, 8 and 11. The parametric estimates have a larger average correlation in each of the simulations. This result is expected since the the true data generating process is linear. As is the case in previous chapter, the estimates given by the nonparametric estimator, although less efficient, are acceptable. Each of the DEA measures however are much less desirable. Many of these conclusions drastically change when the data generating process becomes nonlinear. The results from Tables 6, 9 and 12 show that the nonparametric estimator slightly outperforms the parametric estimator with large N and T. This result is magnified in Tables 7, 10 and 13, when the technology is strictly concave and the parametric estimators do not perform nearly as well as the nonparametric estimators. The NPFGLS estimation exposes the weakness of the parametric estimator when its functional form is slightly misspecified. DEA however has a higher correlation for the case of $\rho = 0.1$ and $T = 5$. One of the main benefits of using DEA is the fact that its production function does not assume a particular functional form and can better adapt to complex technologies. This benefit appears to be nullified with the introduction of NPFGLS and a sufficient size data set. This suggests, that unless the true technology is known *a priori*, that the

---

[5]The results given in the tables are similar to those of Gong and Sickles (1992) between the linear stochastic model and the constant returns to scale DEA model (note that the aforementioned paper uses a much longer time period - from 10 to 50). Therefore the analysis here will focus primarily on the relationship between the parametric and nonparametric stochastic models.

nonparametric kernel estimators of inefficiency be used in practice.

4   Empirical Example

Fare, Grosskopf, Norris and Zhang (1994) employ DEA in order to estimate associated efficiency levels of 17 OECD economies. Their technology contains three macroeconomic variables: aggregate output, labor and physical capital. They let $< Y_t^i, L_t^i, K_t^i >$ represent T observations on these three variables for each of the N countries. The constant-returns-to-scale reference technology (from where the others are obvious) for the 17 countries in period t is defined by

$$P_t = \left\{ < Y, L, K > \in \Re_+^3 \mid Y \leq \sum_j z^i Y_t^i, L \geq \sum_j z^i L_t^i, K \geq \sum_j z^i K_t^i, z^i \geq 0 \ \forall i \right\}, \quad (25)$$

The output based efficiency index for country i at time t is defined by

$$TE_{DEA-CRS} = \min\{\lambda \mid < Y_t^i/\lambda, L_t^i, K_t^i > \in P_t\}. \quad (26)$$

The index is the maximal proportional amount that output can be expanded while remaining technologically feasible, given the technology and input quantities.

 Fare, Grosskopf, Norris and Zhang (1994) use the Penn World Table for data on aggregate output, physical capital and labor. The last three columns of data in Table 10 show the values of technical efficiency for 1983 as determined by each of the DEA procedures.[6] In contrast to stochastic frontiers, DEA typically allows for more than one country to be considered technically efficient (best practice country).

---

[6]The efficiency calculations for DEA are carried out using the software *OnFront (* note that these results differ slightly from the previous paper because of the use of a more recent version of the PWT), available from Economic Measurement and Quality i Lund AB (Box 2134, S-220 02 Lund, Sweden [www.emq.se]).

Those countries with a value of 1.00 (U.K. and U.S.A. for CRS and NRS, whereas VRS adds Ireland and Norway) define the frontier.[7] This leads to the discussion of what returns to scale are appropriate. Most cross-country data research assumes constant returns to scale. However, stochastic tests reject the null hypothesis of CRS in this data set at the 1% level of significance. For purpose of comparison, the stochastic models are estimated with and without the assumption of constant returns to scale.

The first four columns in Table 14 give values of technical efficiency for the parametric and nonparametric models respectively. These are obtained by using 7 periods of data from 1980 to 1986. Reported are the values of time invariant technical efficiency suggested to represent the median year, 1986. As noted previously, stochastic frontiers typically only allow one firm to define the frontier. Under CRS both the parametric and nonparametric procedure designate the U.S.A. as the country that defines the frontier, with mean values of technical efficiency across the 17 countries of 0.75 and 0.73 respectively. These results are very similar to those obtained by DEA under CRS.[8] When the CRS assumption is dropped, the estimates change drastically in some cases. The nonparametric model appears to still work well. Ireland is the country that defines the frontier (DEA with VRS also places Ireland on the frontier) and the U.S.A. is ranked second (Table 15). The parametric model states that Ireland is the country that defines the frontier as well, but places the U.S.A. 16th out of 17 countries in the sample. These results amplify what has been stated in the Monte Carlo exercises. The nonparametric methods dominate when the technology becomes more complex.

---

[7]Figures 1 through 7 show the kernel density estimates of technical efficiency for the each of the procedures. Further, Table 15 shows the ranking from the efficiency scores.

[8]Figures 8-10 plot the production frontiers and production plans.

5 Conclusion

This dissertation proposes nonparametric kernel estimators in order to estimate technical efficiency. This technique allows for more efficient estimation of complex technologies. The monte carlo exercises show that although the parametric model slightly outperforms the nonparametric model when the true technology is linear, when the technology becomes slightly more complex, the nonparametric estimators outperform the now biased parametric results as well as DEA procedures. Thus, it is suggested that the nonparametric method be used in practice because the true technology is generally unknown, and as the Monte Carlo exercises show, the cost of choosing the wrong estimator may be very high.

Appendix

Following Li and Woolridge (2000) we rewrite $d(x)$ as

$$d(x) = \begin{pmatrix} \sum_i K_i w_{ii} G_N \left(\begin{smallmatrix} 1 \\ x_i - x \end{smallmatrix} \right) \left(\begin{smallmatrix} 1 & (x_i - x)' \end{smallmatrix} \right) \\ + \sum_{i \neq j} \sqrt{K_i} \sqrt{K_j} w_{ij} G_N \left(\begin{smallmatrix} 1 \\ x_i - x \end{smallmatrix} \right) \left(\begin{smallmatrix} 1 & (x_i - x)' \end{smallmatrix} \right) \end{pmatrix}^{-1}$$

$$\left(\sum_i K_i w_{ii} G_N \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} y_i + \sum_{i \neq j} \sqrt{K_i} \sqrt{K_j} w_{ij} G_N \begin{pmatrix} 1 \\ x_i - x \end{pmatrix} y_i \right)$$

where

$$G_N = \begin{pmatrix} h^2 & 0 \\ 0 & I_k \end{pmatrix},$$

$$I_k = G_N^{-1} G_N,$$

$$K_i = K \left(\frac{x_i - x}{h} \right)$$

and

$$\Omega_{ij}^{-1} \equiv w_{ij}.$$

Multiplying both the numerator and denominator by $\frac{1}{Nh^{k+2}}$ and substituting the Taylor expansion $\left(y_i = \left(\begin{smallmatrix} 1 & (x_i - x)' \end{smallmatrix} \right) \delta(x) + (x_i - x) \, m''(x) \, (x_i - x)' / 2 + R_m(x_i, x) + \varepsilon_i \right)$ into the above equation gives

$$d(x) = \begin{pmatrix} \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \left(\begin{smallmatrix} h^2 \\ x_i - x \end{smallmatrix} \right) \left(\begin{smallmatrix} 1 & (x_i - x)' \end{smallmatrix} \right) \\ + \frac{1}{Nh^{k+2}} \sum_{i \neq j} \sqrt{K_i} \sqrt{K_j} w_{ij} \left(\begin{smallmatrix} h^2 \\ x_i - x \end{smallmatrix} \right) \left(\begin{smallmatrix} 1 & (x_i - x)' \end{smallmatrix} \right) \end{pmatrix}^{-1}$$

$$\begin{pmatrix} \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \left(\begin{smallmatrix} h \\ x_i - x \end{smallmatrix} \right) \\ \left(\begin{smallmatrix} 1 & (x_i - x)' \end{smallmatrix} \right) \delta(x) \\ + (x_i - x)' m''(x)(x_i - x)/2 + R_m(x_i, x) + \varepsilon_i \end{pmatrix} \\ + \frac{1}{Nh^{k+2}} \sum_{i \neq j} \sqrt{K_i} \sqrt{K_j} w_{ij} \left(\begin{smallmatrix} h^2 \\ x_i - x \end{smallmatrix} \right) \\ \left(\begin{smallmatrix} 1 & (x_i - x)' \end{smallmatrix} \right) \delta(x) \\ + (x_i - x)' m''(x)(x_i - x)/2 + R_m(x_i, x) + \varepsilon_i \end{pmatrix} .$$

After simplifying this expression it becomes obvious that

$$
\begin{aligned}
d(x) &= \delta(x) + \left(\begin{array}{c} \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \begin{pmatrix} h^2 & h^2(x_i - x)' \\ (x_i - x) & (x_i - x)(x_i - x)' \end{pmatrix} \\ + \frac{1}{Nh^{k+2}} \sum_{i\neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} \begin{pmatrix} h^2 & h^2(x_i - x)' \\ (x_i - x) & (x_i - x)(x_i - x)' \end{pmatrix} \end{array} \right)^{-1} \\
&\quad \left(\begin{array}{c} \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \binom{h^2}{x_i - x} \left((x_i - x)' m''(x)(x_i - x)/2 + R_m(x_i, x) + \varepsilon_i \right) \\ + \frac{1}{Nh^{k+2}} \sum_{i\neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} \binom{h^2}{x_i - x} \left((x_i - x)' m''(x)(x_i - x)/2 + R_m(x_i, x) + \varepsilon_i \right) \end{array} \right) \\
&\equiv \delta(x) + \left(A^{1,x} \right)^{-1} \left(A^{2,x} + A^{3,x} \right) + (s.o.)
\end{aligned}
$$

where

$$
A^{1,x} = \begin{array}{c} \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \begin{pmatrix} h^2 & h^2(x_i - x)' \\ (x_i - x) & (x_i - x)(x_i - x)' \end{pmatrix} \\ + \frac{1}{Nh^{k+2}} \sum_{i\neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} \begin{pmatrix} h^2 & h^2(x_i - x)' \\ (x_i - x) & (x_i - x)(x_i - x)' \end{pmatrix} \end{array},
$$

$$
\begin{aligned}
A^{2,x} &= \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \begin{pmatrix} h^2 \\ x_i - x \end{pmatrix} (x_i - x)' m''(x)(x_i - x)/2 \\
&\quad + \frac{1}{Nh^{k+2}} \sum_{i\neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} \begin{pmatrix} h^2 \\ x_i - x \end{pmatrix} (x_i - x)' m''(x)(x_i - x)/2,
\end{aligned}
$$

$$
A^{3,x} = \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \begin{pmatrix} h^2 \\ x_i - x \end{pmatrix} \varepsilon_i + \frac{1}{Nh^{k+2}} \sum_{i\neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} \begin{pmatrix} h^2 \\ x_i - x \end{pmatrix} \varepsilon_i
$$

and

$$
s.o. = \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} \begin{pmatrix} h^2 \\ x_i - x \end{pmatrix} R_m(x_i, x) + \frac{1}{Nh^{k+2}} \sum_{i\neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} \begin{pmatrix} h^2 \\ x_i - x \end{pmatrix} R_m(x_i, x)
$$

where $(s.o.)$ has smaller order than $\left(A^{1,x} \right)^{-1} \left(A^{2,x} \right)$. We can now rewrite the expression as

$$
D(N)(d(x) - \delta(x)) = D(N) \left(A^{1,x} \right)^{-1} \left(A^{2,x} + A^{3,x} \right) + (s.o.).
$$

We will prove the theorem if we prove the following four statements:

(i) $D(N)\left(A^{1,x}\right)^{-1}\left(A^{2,x}+A^{3,x}\right) = D(N)M^{-1}(A^{2,x}+A^{3,x})+o_p(1)$

(ii) $D(N)M^{-1}(A^{2,x}+A^{3,x}) = RD(N)(A^{2,x}+A^{3,x})+o_p(1)$

(iii) $D(N)A^{2,x} = \left(\begin{smallmatrix}(Nh^{k+4})^{\frac{1}{2}}\mu_k f(x)W \\ 0\end{smallmatrix}\right)+o_p(1)$

(iv) $D(N)A^{3,x} \to N(0,V)$ in dist

where

$$M = \left(\begin{array}{cc} f(x)W & 0 \\ c_k f'(x)W & c_k f(x)WI_k \end{array}\right).$$

Further

$$R = diag(M^{-1})$$

and

$$V = \left(\begin{array}{cc} d_k\sigma_\varepsilon^2 f(x)W & 0 \\ 0 & v_k\sigma_\varepsilon^2 f(x)WI_k \end{array}\right).$$

Proof of (i). Note that $D(N)\left(A^{1,x}\right)^{-1}\left(A^{2,x}+A^{3,x}\right) = D(N)M^{-1}(A^{2,x}+A^{3,x})+$ $D(N)\left((A^{1,x})^{-1}-M^{-1}\right)(A^{2,x}+A^{3,x})$. Thus, we only need to show that $D(N)\left((A^{1,x})^{-1}-M^{-1}\right)(A^{2,x}+A^{3,x}) = o_p(1)$. This can be shown by first examining the asymptotic behavior of each element in $A^{1,x}$. Defining

$$A^{1,x} = \left(\begin{array}{cc} A_{11}^{1,x} & A_{12}^{1,x} \\ A_{21}^{1,x} & A_{22}^{1,x} \end{array}\right)$$

yields

$$A_{11}^{1,x} = \frac{1}{Nh^{k+2}}\sum_i K_i w_{ii}h^2 + \frac{1}{Nh^{k+2}}\sum_{i\neq j}\sqrt{K_i}\sqrt{K_j}w_{ij}h^2,$$

$$A_{21}^{1,x} = \frac{1}{Nh^{k+2}}\sum_i K_i w_{ii}(x_i-x) + \frac{1}{Nh^{k+2}}\sum_{i\neq j}\sqrt{K_i}\sqrt{K_j}w_{ij}(x_i-x),$$

50

$$A_{12}^{1,x} = \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} h^2 (x_i - x)' + \frac{1}{Nh^{k+2}} \sum_{i \neq j} \sqrt{K_i} \sqrt{K_j} w_{ij} h^2 (x_i - x)',$$

and

$$A_{22}^{1,x} = \frac{1}{Nh^{k+2}} \sum_i K_i w_{ii} (x_i - x)(x_i - x)' + \frac{1}{Nh^{k+2}} \sum_{i \neq j} \sqrt{K_i} \sqrt{K_j} w_{ij} (x_i - x)(x_i - x)'.$$

By taking the first element we achieve

$$
\begin{aligned}
E(A_{11}^{1,x}) &= \frac{1}{Nh^k} \sum_i E(K_i) w_{ii} + \frac{1}{Nh^k} \sum_{i \neq j} E\left(\sqrt{K_i}\sqrt{K_j}\right) w_{ij} \\
&= \frac{1}{h^k} \int K\left(\frac{x_i - x}{h}\right) f(x_i) dx_i \frac{1}{N} \sum_i w_{ii} \\
&\quad + \frac{1}{h^k} \int K\left(\frac{x_i - x}{h}\right)^{\frac{1}{2}} f(x_i) dx_i \int K\left(\frac{x_j - x}{h}\right)^{\frac{1}{2}} f(x_j) dx_j \frac{1}{N} \sum_{i \neq j} w_{ij} \\
&\to f(x) \frac{1}{N} \sum_i w_{ii} + o_p(1) \\
&= f(x) tr\left(\Omega^{-1}\right) + o_p(1) \\
&= f(x) W + o_p(1).
\end{aligned}
$$

$A_{21}^{1,x}$ is decomposed similarly as

$$
\begin{aligned}
E\left(A_{21}^{1,x}\right) &= \frac{1}{h^2} \int f(x + \psi h) K(\psi) h \psi d\psi \frac{1}{N} \sum_i w_{ii} \\
&\quad + \frac{1}{h^2} \int K\left(\frac{x_i - x}{h}\right)^{\frac{1}{2}} f(x_i) dx_i \int K\left(\frac{x_j - x}{h}\right)^{\frac{1}{2}} f(x_j) dx_j \frac{1}{N} \sum_{i \neq j} w_{ij} \\
&\to c_k f'(x) tr\left(\Omega^{-1}\right) + o_p(1) \\
&= c_k f'(x) W + o_p(1).
\end{aligned}
$$

Next, the term $A_{12}^{1,x} = h^2 (A_{21}^{1,x})' = O_p(h^2)$. Finally, $A_{22}^{1,x}$ can be shown as

$$
\begin{aligned}
E\left(A_{22}^{1,x}\right) &= f(x)\int K(\psi)\psi\psi' d\psi \frac{1}{N}\sum_i w_{ii} \\
&\quad + \frac{1}{h^2}\int K\left(\frac{x_i - x}{h}\right)^{\frac{1}{2}} f(x_i)dx_i \int K\left(\frac{x_j - x}{h}\right)^{\frac{1}{2}} f(x_j)dx_j \frac{1}{N}\sum_{i\neq j} w_{ij} \\
&\rightarrow c_k f(x) tr\left(\Omega^{-1}\right) I_k + o_p(1) \\
&= c_k f(x) W I_k + o_p(1).
\end{aligned}
$$

Thus we have

$$
A^{1,x} = \begin{pmatrix} f(x)W & 0 \\ c_k f'(x)W & c_k f(x)W I_k \end{pmatrix} + o_p(1).
$$

By inverting this matrix (through the method of the partitioned inverse) we achieve

$$
\left(A^{1,x}\right)^{-1} = \begin{pmatrix} \frac{1}{f(x)W} + o_p(1) & O_p(h^2) \\ \frac{-f'(x)}{f^2(x)W} + o_p(1) & \frac{I_k}{c_k f(x)W} + o_p(1) \end{pmatrix}
$$

and thus

$$
\left(A^{1,x}\right)^{-1} - M^{-1} = \begin{pmatrix} o_p(1) & O_p(h^2) \\ o_p(1) & o_p(1)I_k \end{pmatrix}
$$

which completes the proof of (i).

Proof of (ii). This holds because the off diagonal elements of M^{-1} are $o_p(1)$. Specifically,

$$
\begin{aligned}
\left(Nh^{k+2}\right)^{\frac{1}{2}} A_1^{2,x} &= \left(Nh^{k+2}\right)^{\frac{1}{2}} O_p(h^2) \\
&= O_p\left(\left(Nh^{k+6}\right)^{\frac{1}{2}}\right) \\
&= o_p(1)
\end{aligned}
$$

and

$$\left(Nh^{k+2}\right)^{\frac{1}{2}} A_2^{2,x} = \left(Nh^{k+2}\right)^{\frac{1}{2}} O_p\left(\left(Nh^k\right)^{\frac{1}{2}}\right)$$

$$= O_p(h)$$

$$= o_p(1).$$

Proof of (iii). Premultiplying $A^{2,x}$ by $D(N)$ gives

$$D(N) = \begin{pmatrix} \left(Nh^k\right)^{\frac{1}{2}} A_1^{2,x} \\ \left(Nh^{k+2}\right)^{\frac{1}{2}} A_2^{2,x} \end{pmatrix}$$

$$= \begin{pmatrix} \left(Nh^k\right)^{\frac{1}{2}} \begin{pmatrix} \sum_i K_i w_{ii} h^2 (x_i - x)' m''(x)(x_i - x)/2 \\ + \sum_{i \neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} h^2 (x_i - x)' m''(x)(x_i - x)/2 \end{pmatrix} \\ \left(Nh^{k+2}\right)^{\frac{1}{2}} \begin{pmatrix} \sum_i K_i w_{ii} h^2 (x_i - x)' m''(x)(x_i - x)/2 \\ + \sum_{i \neq j} \sqrt{K_i}\sqrt{K_j} w_{ij} h^2 (x_i - x)' m''(x)(x_i - x)/2 \end{pmatrix} \end{pmatrix}$$

$$\rightarrow \begin{pmatrix} \left(Nh^k\right)^{\frac{1}{2}} h^2 \mu_k f(x) W + o_p(1) \\ \left(Nh^{k+2}\right)^{\frac{1}{2}} O_p(h^2) \end{pmatrix}$$

$$= \begin{pmatrix} \left(Nh^k\right)^{\frac{1}{2}} h^2 \mu_k f(x) W \\ 0 \end{pmatrix} + o_p(1)$$

which proves (iii).

Proof of (iv). This proof will examine the variance of each component of $D(N)A^{3,x}$ as well as the covariance between the two components. First,

$$VAR\left(\left(Nh^k\right)^{\frac{1}{2}} A_1^{3,x}\right) = E\left(Nh^k \left(A_1^{3,x}\right)^2\right)$$

$$\rightarrow f(x)\sigma_\varepsilon^2 \int K^2(\psi) d\psi \frac{1}{N} \sum_i w_{ii} + o(1)$$

$$= d_k f(x)\sigma_\varepsilon^2 W + o(1).$$

Next

$$VAR\left((Nh^k)^{\frac{1}{2}}A_2^{3,x}\right) = \sigma_\varepsilon^2 \int f(x+h\psi)K^2(\psi)\psi\psi'd\psi\frac{1}{N}\sum_i w_{ii}$$

$$\rightarrow v_k f(x)\sigma_\varepsilon^2 WI_k + o(1).$$

The covariance is shown to go to

$$COV\left((Nh^k)^{\frac{1}{2}}A_1^{3,x},(Nh^k)^{\frac{1}{2}}A_2^{3,x}\right) = (Nh^{k+1})\,E\left(A_1^{3,x},A_2^{3,x}\right)$$

$$= O(h)$$

$$= o(1).$$

Hence, $VAR\left(D(N)A^{3,x}\right) = V + o(1)$ and since $A^{3,x}$ has a zero mean

$$D(N)A^{3,x} \rightarrow N(0,V).$$

Finally, by proving the four statements, we can show that

$$D(N)\left(d(x)-\delta(x)-\begin{pmatrix}h^2\mu_k W\\0\end{pmatrix}\right) = RD(N)\left(A^{2,x}+A^{3,x}\right)-\begin{pmatrix}h^2\mu_k W\\0\end{pmatrix}+o_p(1)$$

$$= RD(N)A^{3,x}+o_p(1)$$

$$\rightarrow R\left(N(0,V)\right)+o_p(1)$$

$$\rightarrow N(0,RVR)$$

$$= N(0,\Sigma_x).$$

Now we note that $\widehat{\Omega}$ is a consistent estimator of Ω, the proof of it follows from the results of Amemiya (1971) for the parametric model. Thus the asymptotic distribution of $D(N)\left(\widehat{\delta}(x)-\delta(x)\right)$ is the same as that of $D(N)\left(d(x)-\delta(x)\right)$. $\blacksquare$

References

[1] Ahn, S.C., D.H. Good, and R.C. Sickles (2000), "Estimation of Long-Run Inefficiency Levels: A Dynamic Frontier Approach," *Econometric Reviews* 19, 461-92.

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

[3] Amemiya, T. (1971), "The Estimation of the Variances in a Variance-Components Model," *International Economic Review*, 12, 1-13.

[4] Atkinson, S.E., and C. Cornwell (1993), "Measuring Technical Efficiency with Panel Data," *Journal of Econometrics* 6, 257-61.

[5] Baltagi, B.H. (1995), *Econometric Analysis of Panel Data.* New York: Wiley and Sons.

[6] Baltagi, B. H., Y-J. Chang, and Q. Li (1992), "Monte Carlo Results on Several New and Existing Tests for the Error Component Model," *Journal of Econometrics*, 54, 95-120.

[7] Battese, G.E., and T.J. Coelli (1988), "Prediction of Firm-Level Technical Efficiencies with a Generalized Frontier Production Function and Panel Data," *Journal of Econometrics* 38, 387-99.

[8] Battese, G.E., and T.J. Coelli (1992), "Frontier Production Functions, Technical Efficiency and Panel Data: With Application to Paddy Farmers in India," *Journal of Productivity Analysis* 3, 153-69.

[9] Berg, M. D., Q. Li and A. Ullah. (2000), "Instrumental Variable Estimation of Semiparametric Dynamic Panel Data Models: Monte Carlo Results on Several New and Existing Estimators," *Advances in Econometrics*, 15, 297-315.

[10] Carree, M.A. (2002), "Technological Inefficiency and the Skewness of the Error Component in Stochastic Frontier Analysis," *Economics Letters* 77, 101-7.

[11] Caves, D.W., L.R. Christensen, and W.E. Diewert (1982), "The Economic Theory of Index Numbers and the Measurement of Input, Output, and Productivity," *Econometrica*, 50, 1393-414.

[12] Charnes, A., W.W. Cooper, A.Y. Lewin, and L.M. Seiford (1994), *Data Envelopment Analysis: Theory, Methodology, and Application.* Boston: Kluwer Academic Publishers.

[13] Charnes, A., W.W. Cooper, and E. Rhodes (1978), "Measuring the Efficiency of Decision-Making Units," *European Journal of Operational Research* 2, 429-44.

[14] Cornwell, C., P. Schmidt, and R.C. Sickles (1990), "Production Frontiers with Cross-Sectional and Time-Series Variation in Efficiency Levels," *Journal of Econometrics* 46, 185-200.

[15] De Borger, B., and K. Kerstens (1996), "Cost Efficiency of Belgian Local Governments: A Comparative Analysis of FDH, DEA, and Econometric Approaches," *Regional Science and Urban Economics*, 26, 145-70.

[16] Fan, J. and I. Gijbels (1992), "Variable Bandwidth and Local Linear Regression Smoothers," *Annals of Statistics*, 20, 2008-36.

[17] Fare, R., S. Grosskopf, and C.A.K. Lovell (1994), *Production Frontiers*. Cambridge: Cambridge University Press.

[18] Fare, R., S. Grosskopf, M. Norris, and Z. Zhang (1994), "Productivity Growth, Technical Progress, and Efficiency Change in Industralized Countries," *American Economic Review* 84, 66-83.

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

[20] Gong, B.H., and R.C. Sickles (1992), "Finite Sample Evidence on the Performance of Stochastic Frontier Models Using Panel Data," *Journal of Productivity Analysis* 1, 229-61.

[21] Greene, W. H. (2002), *Econometric Analysis*, Upper Saddle River, Prentice Hall.

[22] Hausman, J.A. (1978), "Specification Tests in Econometrics," *Econometrica* 46, 1251-71.

[23] Hausman, J.A., and W.E. Taylor (1981), "Panel Data and Unobservable Individual Effects," *Econometrica* 49, 1377-98.

[24] Henderson, D.J., and A. Ullah (2003), "A Purely Nonparametric Random Effects Estimator," working paper #03-05, University of California at Riverside.

[25] Heshmati, A. (1994), "Estimating Technical Efficiency, Productivity Growth and Selectivity Bias Using Rotating Panel Data: an Application to Swedish Agriculture," Ph.D. Thesis, Department of Economics, Goteborg University.

[26] Heshmati, A. (1998), "Efficiency Measurement in Rotating Panel Data," *Applied Economics* 30, 919-30.

[27] Heshmati, A., and S.C. Kumbhakar (1994), "Farm Heterogeneity and Technical Efficiency: Some Results from Swedish Dairy Farms," *Journal of Productivity Analysis* 5, 45-61.

[28] Hsiao, C., (1986), *Analysis of Panel Data*. Cambridge: Cambridge University Press.

[29] Kim, Y., and P. Schmidt (2000), "A Review and Empirical Comparison of Bayesian and Classical Approaches to Inference on Efficiency Levels in Stochastic Frontier Models with Panel Data," *Journal of Productivity Analysis* 14, 91-118.

[30] Kneip, A., and L. Simar (1996), "A General Framework for Frontier Estimation with Panel Data," *Journal of Productivity Analysis* 7, 187-212.

[31] Kniesner, T. J., and Q. Li. (2002), "Nonlinearity in Dynamic Adjustment: Semiparametric Estimation of Panel Labor Supply," *Empirical Economics*, 27, 131-48.

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

[33] Kumbhakar, S.C. (1987), "The Specification of Technical and Allocative Ineffi- ciency in Stochastic Production and Profit Frontiers," *Journal of Econometrics* 34, 335-48.

[34] Kumbhakar, S.C. (1990), "Production Frontiers, Panel Data, and Time-Varying Technical Inefficiency," *Journal of Econometrics* 46, 201-12.

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

[36] Kumbhakar, S.C. (1992), "Efficiency Estimation Using Rotating Panel Data Models," *Economics Letters* 38, 169-74.

[37] Kumbhakar, S.C. (1997), "Modeling Allocative Inefficiency in a Translog Cost Function and Cost Share Equations: An Exact Relationship," *Journal of Econo- metrics* 76, 351-6.

[38] Kumbhakar, S.C., C.A.K. Lovell (2000), *Stochastic Frontier Analysis.* Cambridge: Cambridge University Press.

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

[40] Li, Q. (1996), "Estimating a Stochastic Production Frontier when the Adjusted Error is Symmetric," *Economics Letters* 52, 221-8.

[41] Li, Q., and J. Wooldridge (2000), "Estimating Semiparametric Economet- rics Models by Local Linear Method: With An Application to Gross-Country Growth," *Annals of Economics and Finance*, 1, 337-57.

[42] Maddala, G. S., and T. D. Mount (1973), "A Comparative Study of Alternative Estimators for Variance components Models Used in Econometric Applications," *Journal of the American Statistical Association,* 68, 324-8.

[43] Meeusen, W., and J. van den Broeck (1977), "Efficiency Estimation from Cobb- Douglas Production Functions with Composed Error," *International Economic Review* 18, 435-44.

[44] Nerlove, M. (1971), "A Note on Error Components Models," *Econometrica,* 39, 383-96.

[45] Pagan, A., and A. Ullah (1999), *Nonparametric Econometrics*, Cambridge, Cam- bridge University Press.

[46] Park, B.U., and L. Simar (1994), "Efficient Semiparametric Estimation in a Sto- chastic Frontier Model," *Journal of the American Statistical Association*, 89, 929-36.

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

[48] Ritter, C. and L. Simar (1997), "Pitfalls of Nornmal-Gamma Stochastic Frontier Models," *Journal of Productivity Analysis*, 8, 167-82.

[49] Robinson, P. M. (1988), "Root-N-Consistent Semiparametric Regression," *Econometrica,* 56, 931-54.

[50] Schmidt, P. (1988), "Estimation of a Fixed-Effect Cobb-Douglas System Using Panel Data," *Journal of Econometrics* 37, 361-80.

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

[52] Simar, L., and P.W. Wilson (2000), "Statistical Inference in Nonparametric Frontier Models: The State of the Art," *Journal of Productivity Analysis* 13, 49-78.

[53] Simar, L., and P.W. Wilson (2001), "A General Methodology for Bootstrapping in Nonparametric Frontier Models," forthcoming *Journal of Applied Statistics*.

[54] Solow, R.W. (1957), "Technical Change and the Aggregate Production Function," *Quarterly Journal of Economics*, 70: 65-94.

[55] Taylor, W. E. (1980), "Small Sample Consideration in Estimation from Panel Data," *Journal of Econometrics,* 13, 203-23.

[56] Tulkens, H. (1986), "La Performance Productive d'un Service Public: Definitions, methodes de mesure, et Application a la Regie des Postes de Belgique," *L'Actualite Economique, Revue d'analyse Economique, Revue d'Analyse Economique* 62, 306-35.

[57] Tulkens, H., and P. Vanden Eeckaut (1995), "Non-parametric Efficiency, Progress and Regress Measures for Panel Data: Methodological Aspects," *European Journal of Operational Research* 80, 474-99.

[58] Vanden Eeckaut, P., H. Tulken, and M-A Jamar (1993), "Cost Efficiency in Belgian Municipalities," in H.O. Fried, C.A.K. Lovell, and S.S. Schmidt, eds., *The Measurement of Productive Efficiency: Techniques and Applications.* New York: Oxford University Press.

Table 1 - Linear Technology ($\gamma = 0$) - Estimates of m

	$T = 3$				$\rho = 0.1$	
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	0.02461	**0.81206**	-0.00294	**0.263450**	0.00101	**0.152292**
$\widehat{m}_{NPFGLS}$	0.03227	1.18472	-0.00918	0.358010	0.02481	0.334819

	$T = 5$				$\rho = 0.1$	
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	-0.00415	**0.62618**	0.00701	**0.262428**	-0.01735	**0.12318**
$\widehat{m}_{NPFGLS}$	0.03685	1.14297	0.00230	0.580311	-0.01315	0.25352

	$T = 3$				$\rho = 0.4$	
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	0.09199	**1.29724**	0.02132	**0.68479**	-0.00176	**0.23748**
$\widehat{m}_{NPFGLS}$	0.04473	1.62509	0.02872	0.95959	0.00232	0.54860

	$T = 5$				$\rho = 0.4$	
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	-0.00965	**1.08460**	0.00787	**0.56886**	0.00764	**0.22496**
$\widehat{m}_{NPFGLS}$	-0.00925	1.62514	0.01551	0.81964	-0.02416	0.34916

	$T = 3$				$\rho = 0.8$	
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	-0.01835	**1.58223**	0.01502	**0.89115**	0.00918	**0.37374**
$\widehat{m}_{NPFGLS}$	0.01257	2.40996	0.04550	1.20262	-0.01792	0.53778

	$T = 5$				$\rho = 0.8$	
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	0.06109	**1.64919**	-0.03037	**0.76996**	0.00668	**0.28034**
$\widehat{m}_{NPFGLS}$	-0.01066	2.40569	-0.01820	1.08072	0.00536	0.44972

Table 2 - Linear Technology ($\gamma = 0$) - Estimates of β

	$T = 3$			$\rho = 0.1$		
	$N = 10$		$N = 20$		$N = 50$	
	$Bias$	MSE	$Bias$	MSE	$Bias$	MSE
$\widehat{\beta}_{FGLS}$	0.04717	**0.43331**	-0.03045	**0.26559**	0.00498	**0.08858**
$\widehat{\beta}_{NPFGLS}$	0.09091	1.30399	-0.02208	0.74755	0.03207	0.54615

	$T = 5$			$\rho = 0.1$		
	$N = 10$		$N = 20$		$N = 50$	
	$Bias$	MSE	$Bias$	MSE	$Bias$	MSE
$\widehat{\beta}_{FGLS}$	-0.00855	**0.25194**	0.01213	**0.17773**	-0.00236	**0.04905**
$\widehat{\beta}_{NPFGLS}$	0.06144	2.61958	-0.00354	1.48246	-0.00432	0.74635

	$T = 3$			$\rho = 0.4$		
	$N = 10$		$N = 20$		$N = 50$	
	$Bias$	MSE	$Bias$	MSE	$Bias$	MSE
$\widehat{\beta}_{FGLS}$	0.03625	**0.46752**	-0.00015	**0.18585**	-0.00420	**0.06197**
$\widehat{\beta}_{NPFGLS}$	-0.03703	1.32680	-0.02840	0.79431	0.00041	0.48636

	$T = 5$			$\rho = 0.4$		
	$N = 10$		$N = 20$		$N = 50$	
	$Bias$	MSE	$Bias$	MSE	$Bias$	MSE
$\widehat{\beta}_{FGLS}$	0.00386	**0.17615**	0.01756	**0.08143**	0.00871	**0.03467**
$\widehat{\beta}_{NPFGLS}$	0.00991	1.96978	-0.00893	1.08968	-0.04652	0.55036

	$T = 3$			$\rho = 0.8$		
	$N = 10$		$N = 20$		$N = 50$	
	$Bias$	MSE	$Bias$	MSE	$Bias$	MSE
$\widehat{\beta}_{FGLS}$	0.02128	**0.20419**	-0.01315	**0.08525**	-0.01586	**0.03017**
$\widehat{\beta}_{NPFGLS}$	0.01243	0.77846	-0.00564	0.43128	-0.03533	0.23332

	$T = 5$			$\rho = 0.8$		
	$N = 10$		$N = 20$		$N = 50$	
	$Bias$	MSE	$Bias$	MSE	$Bias$	MSE
$\widehat{\beta}_{FGLS}$	0.00467	**0.05750**	-0.00071	**0.02854**	0.00137	**0.01069**
$\widehat{\beta}_{NPFGLS}$	-0.04696	0.92974	0.01424	0.52818	-0.00487	0.24382

Table 3 - Quadratic Technology ($\gamma = 2$) - Estimates of m

	$T = 3$			$\rho = 0.1$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	3.75215	13.95897	3.78125	14.36996	1.38247	13.83682
$\widehat{m}_{NPFGLS}$	1.25748	**2.259489**	1.01324	**1.487497**	-0.64823	**0.69626**

	$T = 5$			$\rho = 0.1$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	3.19402	11.45857	3.21542	11.81433	3.27112	11.80253
$\widehat{m}_{NPFGLS}$	0.60992	**1.52437**	0.45886	**0.80338**	0.329321	**0.42009**

	$T = 3$			$\rho = 0.4$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	3.59057	13.24741	3.73497	14.14550	3.69340	13.72657
$\widehat{m}_{NPFGLS}$	1.45033	**2.18517**	0.85695	**1.08555**	0.60328	**0.67794**

	$T = 5$			$\rho = 0.4$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	3.21889	12.00314	3.28223	11.51401	3.30367	11.18568
$\widehat{m}_{NPFGLS}$	0.70980	**1.92347**	0.43402	**0.98756**	0.25790	**0.41860**

	$T = 3$			$\rho = 0.8$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	3.75286	14.88077	-2.35830	14.53756	3.67594	14.25223
$\widehat{m}_{NPFGLS}$	0.96821	**2.66539**	0.75107	**1.43185**	0.44664	**0.64704**

	$T = 5$			$\rho = 0.8$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{m}_{FGLS}$	3.39286	13.66158	3.27061	11.75863	3.25525	11.04261
$\widehat{m}_{NPFGLS}$	0.54470	**2.216837**	0.26445	**1.12261**	0.13944	**0.45376**

Table 4 - Quadratic Technology ($\gamma = 2$) - Estimates of β

	$T = 3$			$\rho = 0.1$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{\beta}_{FGLS}$	2.76221	8.66174	2.90400	8.97976	2.89316	8.53274
$\widehat{\beta}_{NPFGLS}$	-0.26873	**1.69491**	-0.26373	**0.85174**	-0.36253	**0.53852**

	$T = 5$			$\rho = 0.1$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{\beta}_{FGLS}$	5.87036	35.40215	5.93808	35.62152	6.05484	37.40021
$\widehat{\beta}_{NPFGLS}$	0.12010	**2.86717**	-0.23052	**1.61403**	-0.51341	**1.21027**

	$T = 3$			$\rho = 0.4$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{\beta}_{FGLS}$	3.03288	10.38349	3.06586	9.88796	3.12776	10.00722
$\widehat{\beta}_{NPFGLS}$	-0.05755	**1.43259**	-0.30976	**0.88407**	-0.28894	**0.48019**

	$T = 5$			$\rho = 0.4$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{\beta}_{FGLS}$	5.68197	33.21481	5.92401	35.38354	6.02191	36.17050
$\widehat{\beta}_{NPFGLS}$	0.22978	**2.18748**	-0.24992	**1.28177**	-0.54674	**0.96792**

	$T = 3$			$\rho = 0.8$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{\beta}_{FGLS}$	3.31846	12.05829	3.44486	12.29401	3.51921	12.52890
$\widehat{\beta}_{NPFGLS}$	0.03322	**0.98089**	-0.08389	**0.51563**	-0.12502	**0.25223**

	$T = 5$			$\rho = 0.8$		
	$N = 10$		$N = 20$		$N = 50$	
	Bias	MSE	Bias	MSE	Bias	MSE
$\widehat{\beta}_{FGLS}$	5.79115	34.09982	5.91963	35.26351	5.96645	35.51767
$\widehat{\beta}_{NPFGLS}$	0.09283	**0.96495**	-0.23768	**0.72277**	-0.51123	**0.57375**

Table 5 - ($\alpha = 1$) - Half-Normal Distribution

$\rho = 0.1$

| | $T = 5$ | | | | $T = 7$ | | |
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.6558**	**0.6297**	**0.6204**		**0.7230**	**0.7450**	**0.6776**
TE_{NP}	0.6324	0.6163	0.6194		0.7192	0.7407	0.6774
$TE_{DEA-CRS}$	0.4091	0.4185	0.4163		0.4550	0.5378	0.5411
$TE_{DEA-NRS}$	0.4167	0.4268	0.4295		0.4673	0.5480	0.5241
$TE_{DEA-VRS}$	0.4329	0.4382	0.4301		0.4490	0.5741	0.5863

$\rho = 0.4$

| | $T = 5$ | | | | $T = 7$ | | |
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.6532**	**0.7755**	**0.8629**		**0.7601**	**0.9141**	**0.7646**
TE_{NP}	0.6458	0.7708	0.8423		0.7565	0.9122	0.7636
$TE_{DEA-CRS}$	0.4400	0.4994	0.5938		0.6294	0.6523	0.6721
$TE_{DEA-NRS}$	0.4453	0.5148	0.5944		0.6347	0.6549	0.6756
$TE_{DEA-VRS}$	0.4150	0.5280	0.6202		0.6234	0.6661	0.6841

$\rho = 0.8$

| | $T = 5$ | | | | $T = 7$ | | |
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.8033**	**0.8435**	**0.9694**		**0.9684**	**0.9863**	**0.9833**
TE_{NP}	0.7867	0.8290	0.9622		0.9651	0.9853	0.9822
$TE_{DEA-CRS}$	0.6107	0.7179	0.7303		0.6406	0.7438	0.8438
$TE_{DEA-NRS}$	0.6190	0.7200	0.7298		0.6419	0.7446	0.8446
$TE_{DEA-VRS}$	0.6292	0.7252	0.7299		0.6303	0.7507	0.8507

Table 6 - $(\alpha = \frac{1}{2})$ - Half-Normal Distribution

$\rho = 0.1$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.5427**	**0.5289**	**0.5933**		**0.7448**	**0.7708**	0.7834
TE_{NP}	0.5326	0.5261	0.5950		0.7482	0.7663	**0.7872**
$TE_{DEA-CRS}$	0.3906	0.2619	0.3477		0.4201	0.4801	0.5213
$TE_{DEA-NRS}$	0.3965	0.2744	0.3633		0.4323	0.4863	0.5248
$TE_{DEA-VRS}$	0.3919	0.2801	0.3581		0.4628	0.4849	0.5232

$\rho = 0.4$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.7703**	**0.8588**	**0.8923**		**0.8338**	**0.9168**	0.9207
TE_{NP}	0.7686	0.8571	0.8914		0.8331	0.9278	**0.9233**
$TE_{DEA-CRS}$	0.5253	0.5396	0.5981		0.5475	0.6373	0.6512
$TE_{DEA-NRS}$	0.5299	0.5447	0.6079		0.5579	0.6355	0.6501
$TE_{DEA-VRS}$	0.5352	0.5602	0.6089		0.5593	0.6245	0.6488

$\rho = 0.8$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.9773**	**0.9757**	**0.9799**		**0.9753**	0.9811	0.9890
TE_{NP}	0.9767	0.9749	0.9795		0.9737	**0.9907**	**0.9896**
$TE_{DEA-CRS}$	0.5419	0.6773	0.6947		0.6115	0.6490	0.6512
$TE_{DEA-NRS}$	0.5484	0.6890	0.6975		0.6281	0.6486	0.6502
$TE_{DEA-VRS}$	0.5472	0.6767	0.6940		0.6328	0.6478	0.6498

Table 7 - $(\alpha = 2)$ - Half-Normal Distribution

$\rho = 0.1$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.3076	0.3345	0.3480		0.3849	0.4248	0.4760
TE_{NP}	0.3616	0.3956	0.4275		**0.4352**	**0.4689**	**0.5158**
$TE_{DEA-CRS}$	0.3440	0.3721	0.4065		0.3885	0.3970	0.4002
$TE_{DEA-NRS}$	0.3383	0.3665	0.3918		0.3801	0.3891	0.3982
$TE_{DEA-VRS}$	**0.4127**	**0.4347**	**0.4401**		0.4209	0.4121	0.4310

$\rho = 0.4$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.3818	0.4722	0.6233		0.5215	0.7635	0.7790
TE_{NP}	**0.4885**	**0.5565**	**0.6921**		**0.5843**	**0.8069**	**0.8125**
$TE_{DEA-CRS}$	0.3722	0.4229	0.4637		0.4764	0.5943	0.6055
$TE_{DEA-NRS}$	0.3625	0.4142	0.4549		0.4632	0.5829	0.5952
$TE_{DEA-VRS}$	0.4675	0.5077	0.5141		0.5484	0.6327	0.6421

$\rho = 0.8$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.6020	0.6832	0.6966		0.7276	0.8512	0.9121
TE_{NP}	**0.7181**	**0.7577**	**0.7825**		**0.7876**	**0.8721**	**0.9357**
$TE_{DEA-CRS}$	0.5273	0.5720	0.5992		0.5975	0.6175	0.6235
$TE_{DEA-NRS}$	0.5152	0.5476	0.5844		0.5816	0.6013	0.6185
$TE_{DEA-VRS}$	0.5980	0.6026	0.6823		0.6461	0.6870	0.6987

Table 8 - ($\alpha = 1$) - Exponential Distribution

$\rho = 0.1$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.5666**	**0.5732**	**0.6427**		**0.7267**	**0.8457**	**0.8755**
TE_{NP}	0.5548	0.5698	0.6305		0.7213	0.8436	0.8751
$TE_{DEA-CRS}$	0.3598	0.4005	0.4404		0.3516	0.4632	0.4932
$TE_{DEA-NRS}$	0.3642	0.4046	0.4546		0.3616	0.4819	0.4952
$TE_{DEA-VRS}$	0.3313	0.4027	0.4483		0.3757	0.4876	0.4988

$\rho = 0.4$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.8676**	**0.8718**	**0.9039**		**0.8762**	**0.9116**	**0.9122**
TE_{NP}	0.8547	0.8717	0.8884		0.8724	0.9080	0.9099
$TE_{DEA-CRS}$	0.5471	0.5500	0.5617		0.5060	0.5648	0.5689
$TE_{DEA-NRS}$	0.5553	0.5603	0.5722		0.5079	0.5752	0.5832
$TE_{DEA-VRS}$	0.5465	0.5646	0.5998		0.5269	0.6084	0.6121

$\rho = 0.8$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.9732**	**0.9734**	**0.9829**		**0.9582**	**0.9755**	**0.9812**
TE_{NP}	0.9625	0.9632	0.9736		0.9543	0.9729	0.9802
$TE_{DEA-CRS}$	0.6733	0.7144	0.7703		0.7129	0.8016	0.8214
$TE_{DEA-NRS}$	0.6725	0.7200	0.7729		0.7171	0.8147	0.8222
$TE_{DEA-VRS}$	0.6681	0.7128	0.7794		0.7038	0.8049	0.8217

Table 9 - ($\alpha = \frac{1}{2}$) - Exponential Distribution

$\rho = 0.1$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.6009**	**0.6133**	0.7202		**0.6240**	**0.7067**	0.7512
TE_{NP}	0.5969	0.6088	**0.7265**		0.6232	0.7026	**0.7518**
$TE_{DEA-CRS}$	0.3110	0.3243	0.4776		0.2483	0.3939	0.4852
$TE_{DEA-NRS}$	0.3257	0.3407	0.4904		0.2496	0.3989	0.5011
$TE_{DEA-VRS}$	0.3139	0.3355	0.4957		0.2275	0.4089	0.5050

$\rho = 0.4$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.8566**	**0.8902**	**0.9108**		**0.7537**	**0.7840**	0.8120
TE_{NP}	0.8562	0.8895	0.9063		0.7460	0.7767	**0.8220**
$TE_{DEA-CRS}$	0.4444	0.5278	0.5581		0.4585	0.4633	0.4921
$TE_{DEA-NRS}$	0.4494	0.5370	0.5747		0.4666	0.4769	0.5014
$TE_{DEA-VRS}$	0.4533	0.5519	0.5854		0.4583	0.4646	0.5144

$\rho = 0.8$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.9543**	**0.9830**	**0.9890**		**0.9780**	**0.9849**	0.9925
TE_{NP}	0.9528	0.9823	0.9885		0.9776	0.9841	**0.9929**
$TE_{DEA-CRS}$	0.5405	0.6605	0.6684		0.5442	0.5541	0.5899
$TE_{DEA-NRS}$	0.5487	0.6594	0.6696		0.5463	0.5727	0.6046
$TE_{DEA-VRS}$	0.5788	0.6735	0.6930		0.5421	0.5630	0.6046

Table 10 - ($\alpha = 2$) - Exponential Distribution

$\rho = 0.1$

| | $T = 5$ | | | | $T = 7$ | | |
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.2435	0.2800	0.2918		0.2412	0.5105	0.6214
TE_{NP}	0.2868	0.3428	0.3650		**0.2855**	**0.5539**	**0.6521**
$TE_{DEA-CRS}$	0.2459	0.3277	0.3447		0.3136	0.4485	0.4821
$TE_{DEA-NRS}$	0.2394	0.3177	0.3371		0.2975	0.4421	0.4800
$TE_{DEA-VRS}$	**0.2929**	**0.3531**	**0.3609**		0.3735	0.4952	0.5014

$\rho = 0.4$

| | $T = 5$ | | | | $T = 7$ | | |
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.4316	0.4500	0.5781		0.3518	0.6269	0.7105
TE_{NP}	**0.5447**	**0.5461**	**0.6522**		**0.4193**	**0.6808**	**0.7521**
$TE_{DEA-CRS}$	0.3935	0.4356	0.4435		0.4160	0.5563	0.5744
$TE_{DEA-NRS}$	0.3766	0.4255	0.4301		0.3938	0.5423	0.5698
$TE_{DEA-VRS}$	0.4515	0.5088	0.5196		0.4530	0.6145	0.6345

$\rho = 0.8$

| | $T = 5$ | | | | $T = 7$ | | |
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.6263	0.7106	0.7500		0.7349	0.8945	0.9314
TE_{NP}	**0.7087**	**0.8118**	**0.8243**		**0.7823**	**0.9217**	**0.9641**
$TE_{DEA-CRS}$	0.4862	0.5353	0.5614		0.5647	0.6306	0.6791
$TE_{DEA-NRS}$	0.4791	0.5259	0.5478		0.5482	0.6187	0.6682
$TE_{DEA-VRS}$	0.5878	0.6287	0.6975		0.6686	0.7402	0.7716

Table 11 - ($\alpha = 1$) - Gamma Distribution

$\rho = 0.1$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.5784**	**0.6088**	**0.7080**		**0.3698**	**0.7062**	**0.8063**
TE_{NP}	0.5592	0.6018	0.6874		0.3641	0.7047	0.8053
$TE_{DEA-CRS}$	0.4050	0.4132	0.4300		0.3209	0.4127	0.4844
$TE_{DEA-NRS}$	0.4115	0.4279	0.4407		0.3293	0.4142	0.4922
$TE_{DEA-VRS}$	0.3898	0.4172	0.4501		0.3107	0.4193	0.5011

$\rho = 0.4$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.7488**	**0.7912**	**0.8783**		**0.8120**	**0.8589**	**0.9001**
TE_{NP}	0.7437	0.7908	0.8639		0.8067	0.8529	0.8923
$TE_{DEA-CRS}$	0.3970	0.4212	0.6242		0.7651	0.7860	0.8112
$TE_{DEA-NRS}$	0.3980	0.4212	0.6242		0.7542	0.7842	0.8080
$TE_{DEA-VRS}$	0.4004	0.4992	0.6100		0.7415	0.7723	0.7989

$\rho = 0.8$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.9754**	**0.9776**	**0.9851**		**0.9680**	**0.9880**	**0.9912**
TE_{NP}	0.9653	0.9679	0.9790		0.9656	0.9870	0.9905
$TE_{DEA-CRS}$	0.6120	0.6212	0.6321		0.7951	0.8210	0.8311
$TE_{DEA-NRS}$	0.6128	0.6258	0.6332		0.7911	0.8141	0.8249
$TE_{DEA-VRS}$	0.6241	0.6421	0.6574		0.7895	0.8010	0.8241

Table 12 - $(\alpha = \frac{1}{2})$ - Gamma Distribution

$\rho = 0.1$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.2934**	**0.6117**	**0.6762**		**0.5616**	**0.6437**	0.7109
TE_{NP}	0.2901	0.6092	0.6567		0.5597	0.6405	**0.7122**
$TE_{DEA-CRS}$	0.1889	0.3233	0.4166		0.3372	0.3511	0.3755
$TE_{DEA-NRS}$	0.2000	0.3226	0.4217		0.3502	0.3554	0.3787
$TE_{DEA-VRS}$	0.1946	0.3514	0.4480		0.3663	0.3467	0.3689

$\rho = 0.4$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.6807**	**0.8698**	**0.9101**		**0.8606**	**0.9114**	0.9222
TE_{NP}	0.6741	0.8689	0.9078		0.8604	0.9109	**0.9228**
$TE_{DEA-CRS}$	0.4021	0.4133	0.4880		0.5041	0.5123	0.5244
$TE_{DEA-NRS}$	0.4032	0.4137	0.4880		0.5089	0.5178	0.5268
$TE_{DEA-VRS}$	0.3912	0.4005	0.4198		0.5011	0.5089	0.5189

$\rho = 0.8$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	**0.9787**	**0.9804**	**0.9851**		**0.9804**	**0.9850**	0.9861
TE_{NP}	0.9776	0.9704	0.9802		0.9794	0.9838	**0.9882**
$TE_{DEA-CRS}$	0.6521	0.6711	0.6921		0.7120	0.7210	0.7318
$TE_{DEA-NRS}$	0.6514	0.6741	0.6974		0.7141	0.7211	0.7344
$TE_{DEA-VRS}$	0.6321	0.6521	0.6751		0.7021	0.7052	0.7177

Table 13 - ($\alpha = 2$) - Gamma Distribution

$\rho = 0.1$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.1836	0.2939	0.3256		0.4081	0.4835	0.5546
TE_{NP}	0.2505	0.3527	0.4157		**0.4634**	**0.5321**	**0.5988**
$TE_{DEA-CRS}$	0.2227	0.3297	0.3354		0.4211	0.4241	0.4320
$TE_{DEA-NRS}$	0.2096	0.3224	0.3325		0.4086	0.4091	0.4109
$TE_{DEA-VRS}$	**0.2766**	**0.3645**	**0.3741**		0.4471	0.4777	0.4977

$\rho = 0.4$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.4151	0.4907	0.5765		0.6330	0.6396	0.6522
TE_{NP}	**0.5189**	**0.5775**	**0.6450**		**0.6924**	**0.7003**	**0.7249**
$TE_{DEA-CRS}$	0.3174	0.3205	0.3648		0.3789	0.4790	0.5000
$TE_{DEA-NRS}$	0.3152	0.3127	0.3572		0.3651	0.4629	0.4973
$TE_{DEA-VRS}$	0.3573	0.3814	0.4246		0.4210	0.5074	0.5318

$\rho = 0.8$

	$T = 5$				$T = 7$		
	$N = 10$	$N = 20$	$N = 50$		$N = 10$	$N = 20$	$N = 50$
TE_P	0.6459	0.6702	0.7566		0.8239	0.8300	0.8865
TE_{NP}	**0.7553**	**0.7610**	**0.8195**		**0.8657**	**0.8687**	**0.8978**
$TE_{DEA-CRS}$	0.4121	0.4409	0.3597		0.5256	0.7309	0.7416
$TE_{DEA-NRS}$	0.4112	0.4377	0.3505		0.5123	0.7245	0.7349
$TE_{DEA-VRS}$	0.4154	0.4165	0.3301		0.6100	0.7411	0.7488

Table 14 - Efficiency Values

Country	P	NP	$P{-}CRS$	$NP{-}CRS$	$DEA{-}CRS$	$DEA{-}NRS$	$DEA{-}VRS$
AUSTRALIA	0.56	0.47	0.81	0.81	0.88	0.88	0.91
AUSTRIA	0.70	0.45	0.74	0.72	0.75	0.75	0.86
BELGIUM	0.72	0.45	0.79	0.80	0.85	0.85	0.94
CANADA	0.47	0.54	0.86	0.86	0.90	0.90	0.93
DENMARK	0.74	0.50	0.70	0.70	0.70	0.70	0.85
FINLAND	0.73	0.46	0.63	0.58	0.73	0.73	0.86
FRANCE	0.31	0.52	0.81	0.82	0.85	0.85	0.86
GERMANY, W	0.25	0.40	0.66	0.58	0.84	0.84	0.85
GREECE	0.50	0.35	0.57	0.55	0.61	0.61	0.66
IRELAND	1.00	1.00	0.70	0.70	0.76	0.76	1.00
ITALY	0.34	0.56	0.82	0.82	0.86	0.86	0.88
JAPAN	0.13	0.39	0.55	0.55	0.60	0.61	0.61
NORWAY	0.89	0.56	0.59	0.59	0.83	0.83	1.00
SPAIN	0.37	0.46	0.75	0.75	0.84	0.84	0.86
SWEDEN	0.68	0.44	0.79	0.79	0.80	0.80	0.89
U.K.	0.30	0.61	0.85	0.85	1.00	1.00	1.00
U.S.A.	0.19	0.71	1.00	1.00	1.00	1.00	1.00
Mean	**0.52**	**0.52**	**0.75**	**0.73**	**0.81**	**0.81**	**0.88**

Table 15 - Rankings from Efficiency Scores

Country	P	NP	$P-CRS$	$NP-CRS$	$DEA-CRS$	$DEA-NRS$	$DEA-VRS$
AUSTRALIA	8	9	5	6	4	4	7
AUSTRIA	6	13	10	10	13	13	10
BELGIUM	5	12	7	7	6	6	5
CANADA	10	6	3	2	3	3	6
DENMARK	3	8	12	11	15	15	15
FINLAND	4	11	15	14	14	14	12
FRANCE	13	7	6	4	7	7	11
GERMANY, W	15	15	14	15	8	8	14
GREECE	9	17	17	17	16	16	16
IRELAND	1	1	11	12	12	12	1
ITALY	12	5	4	5	5	5	9
JAPAN	17	16	16	16	17	17	17
NORWAY	2	4	13	13	10	10	1
SPAIN	11	10	9	9	9	9	13
SWEDEN	7	14	8	8	11	11	8
U.K.	14	3	2	3	1	1	1
U.S.A.	16	2	1	1	1	1	1

Figure 1

1983 Distribution of the Parametric Efficiency Index
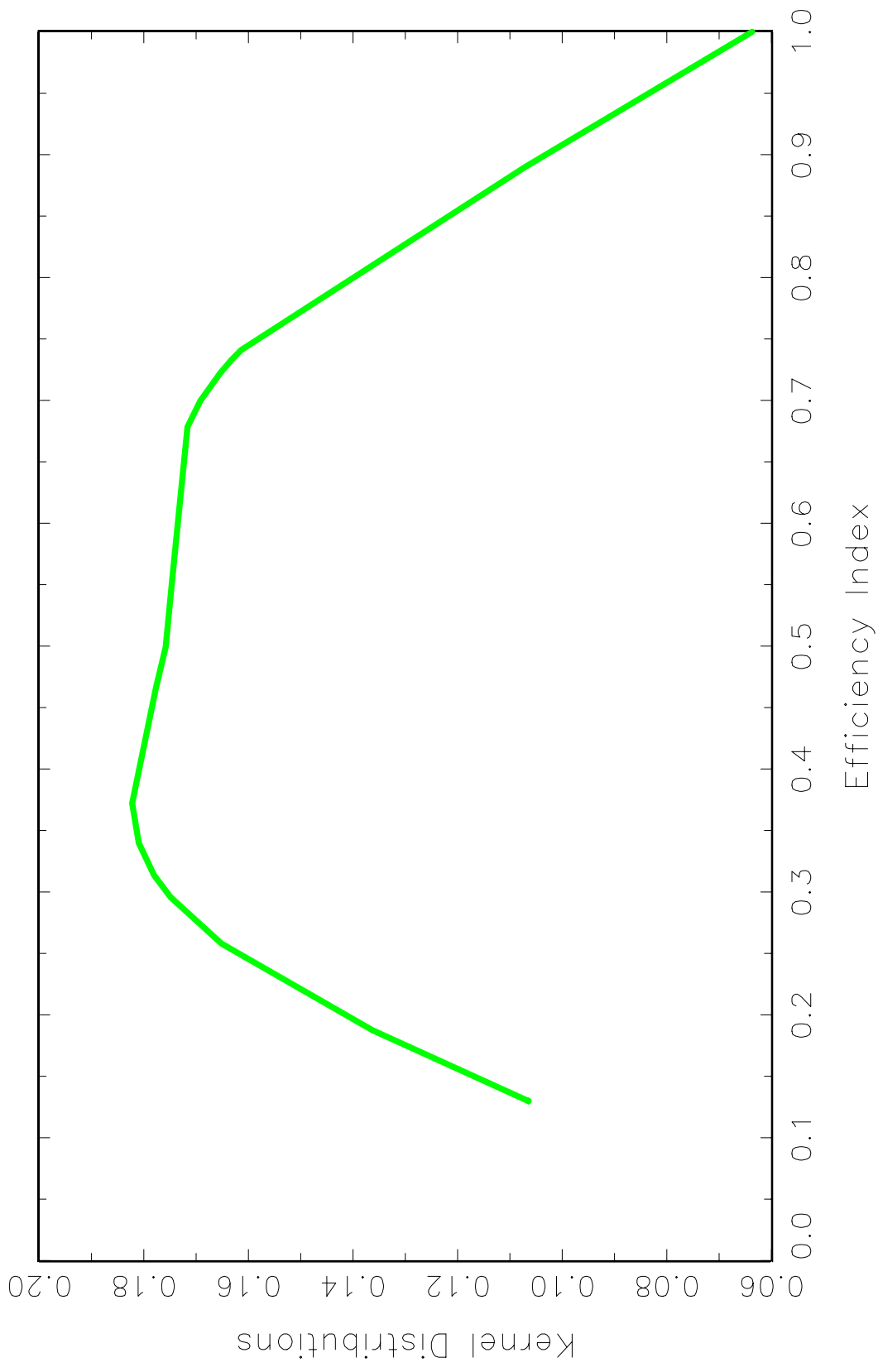
Figure 2

1983 Distribution of the Nonparametric Efficiency Index
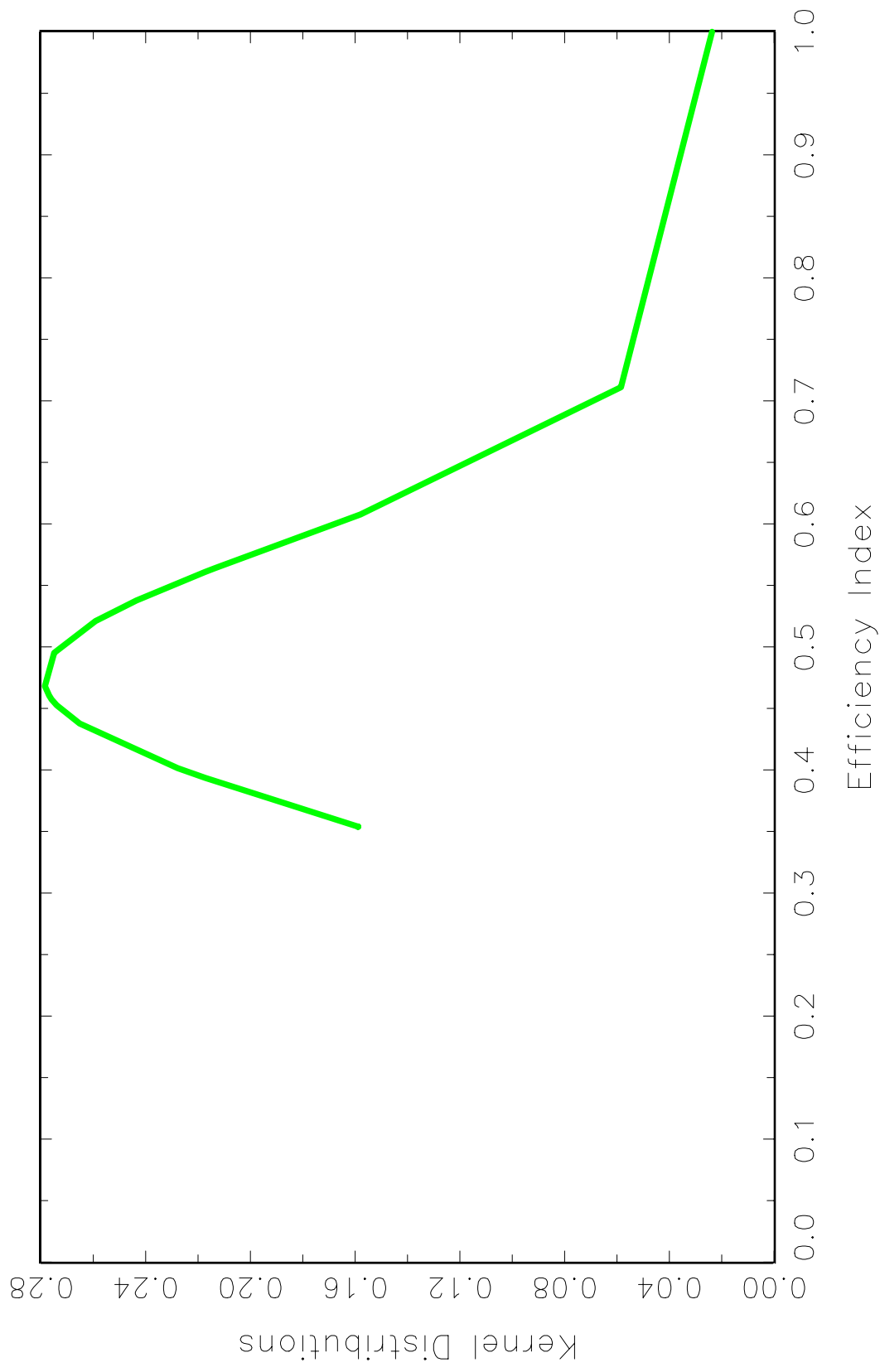
Figure 3

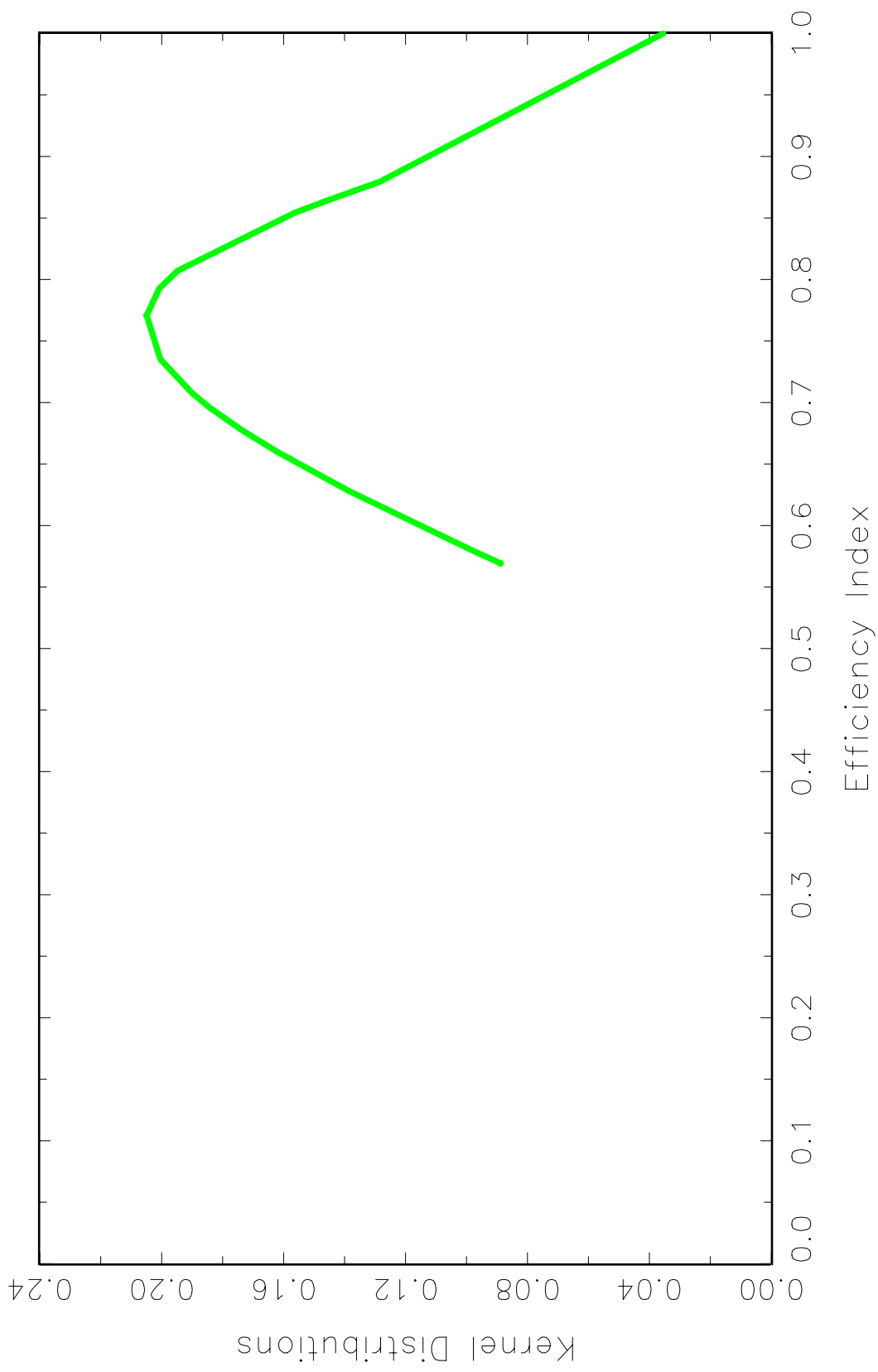1983 Distribution of the Parametric-CRS Efficiency Index

Figure 4

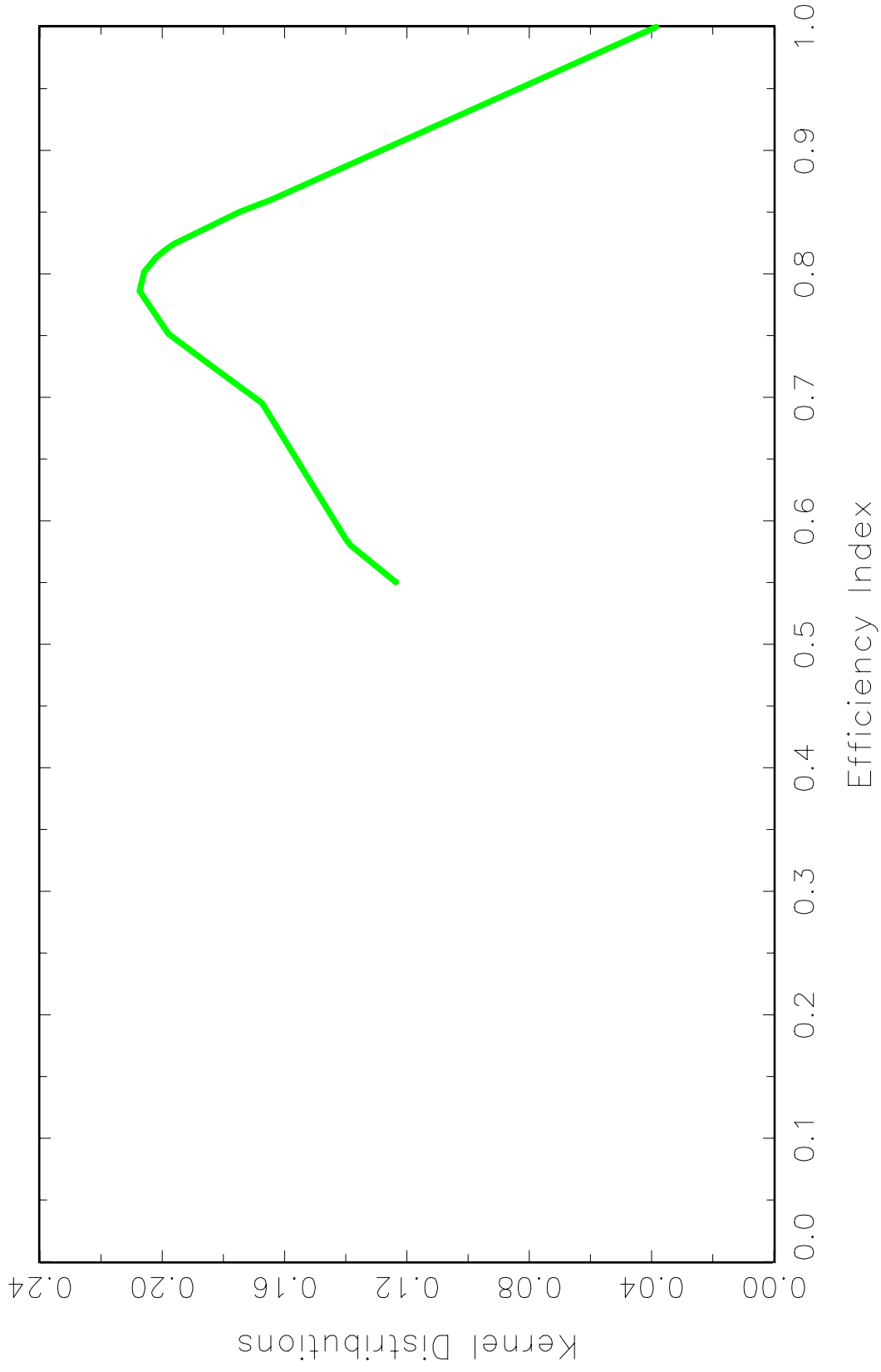1983 Distribution of the Nonparametric-CRS Efficiency Index

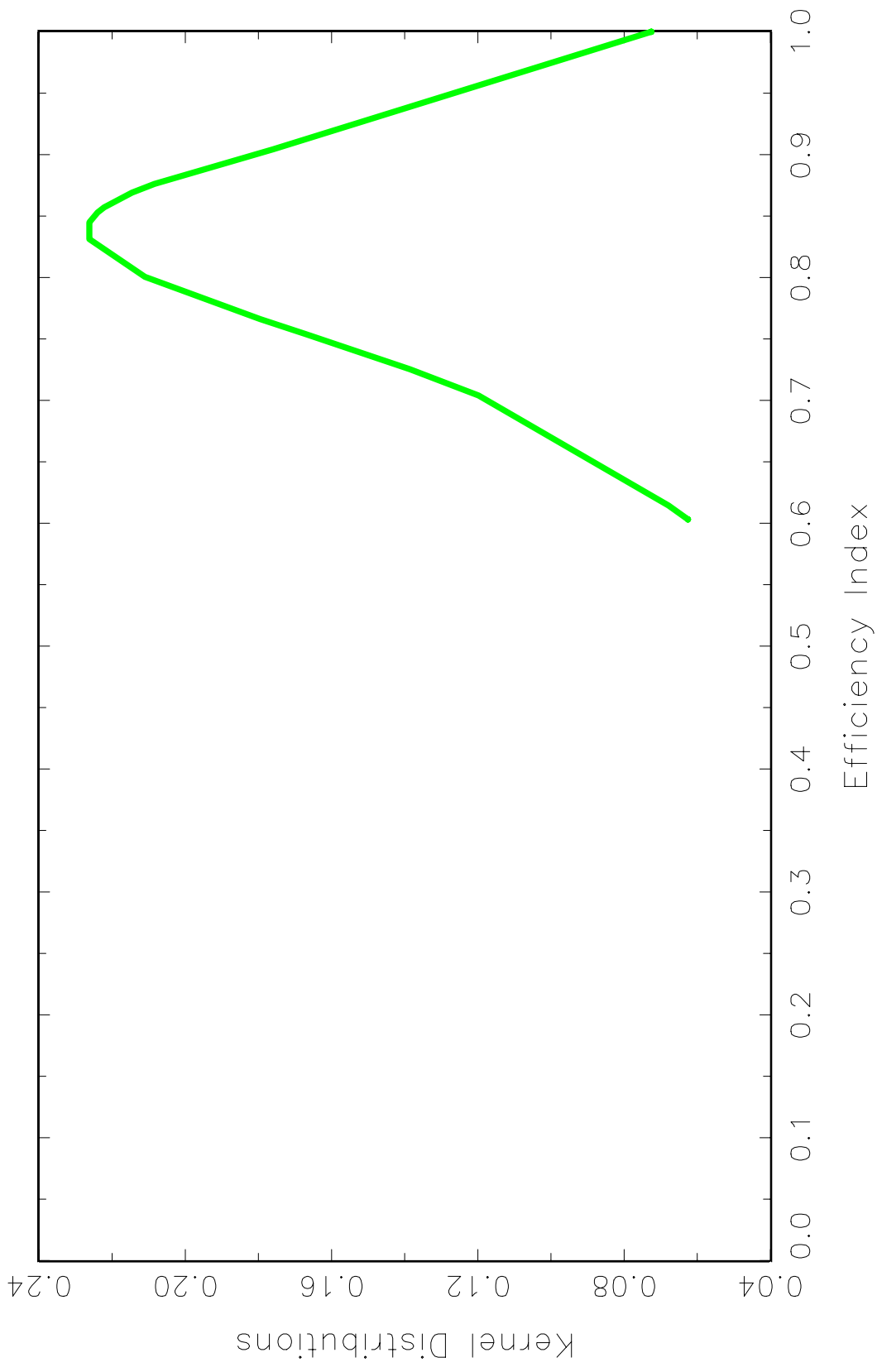Figure 5

1983 Distribution of the DEA–CRS Efficiency Index

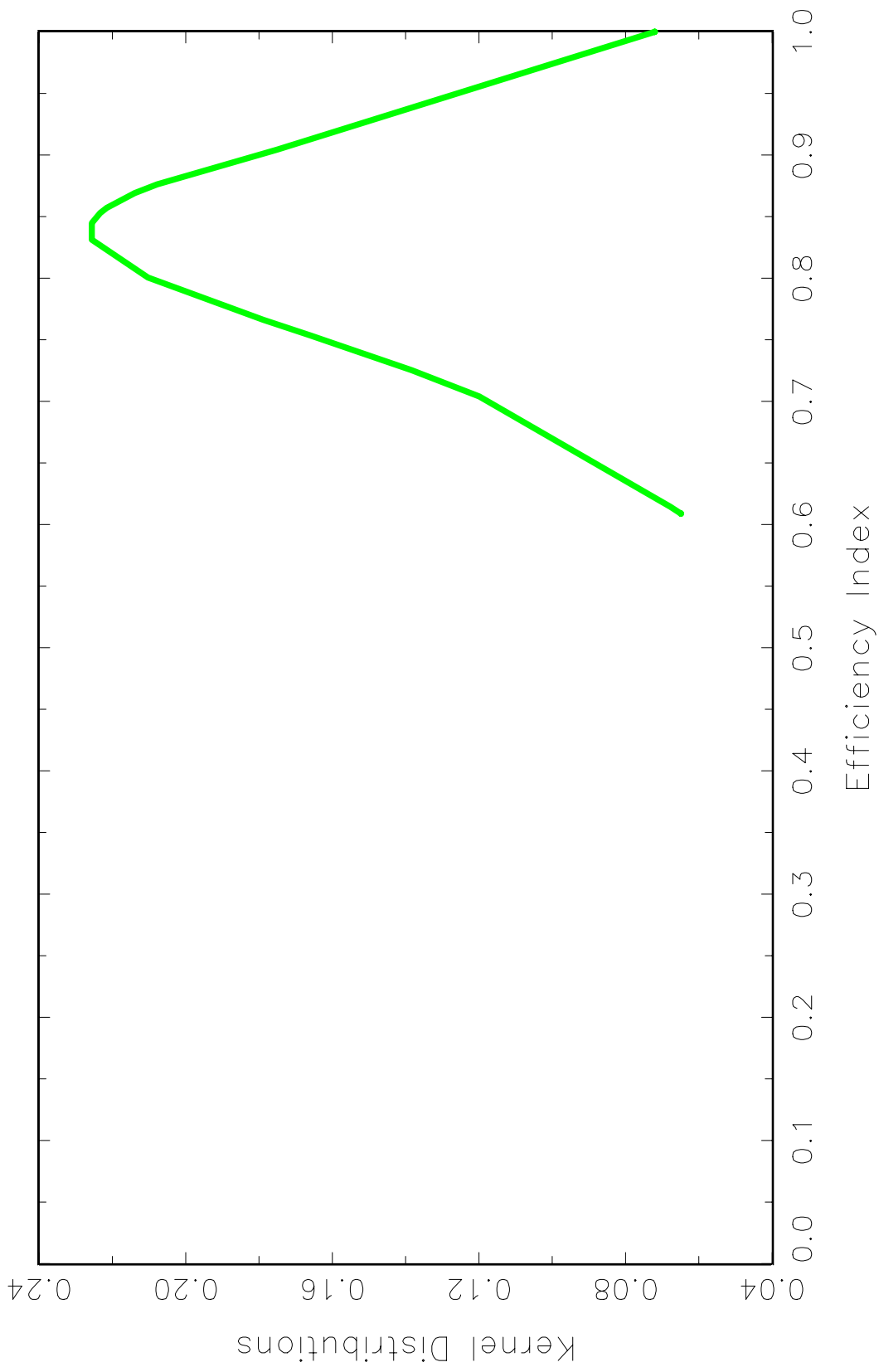Figure 6
1983 Distribution of the DEA-NRS Efficiency Index

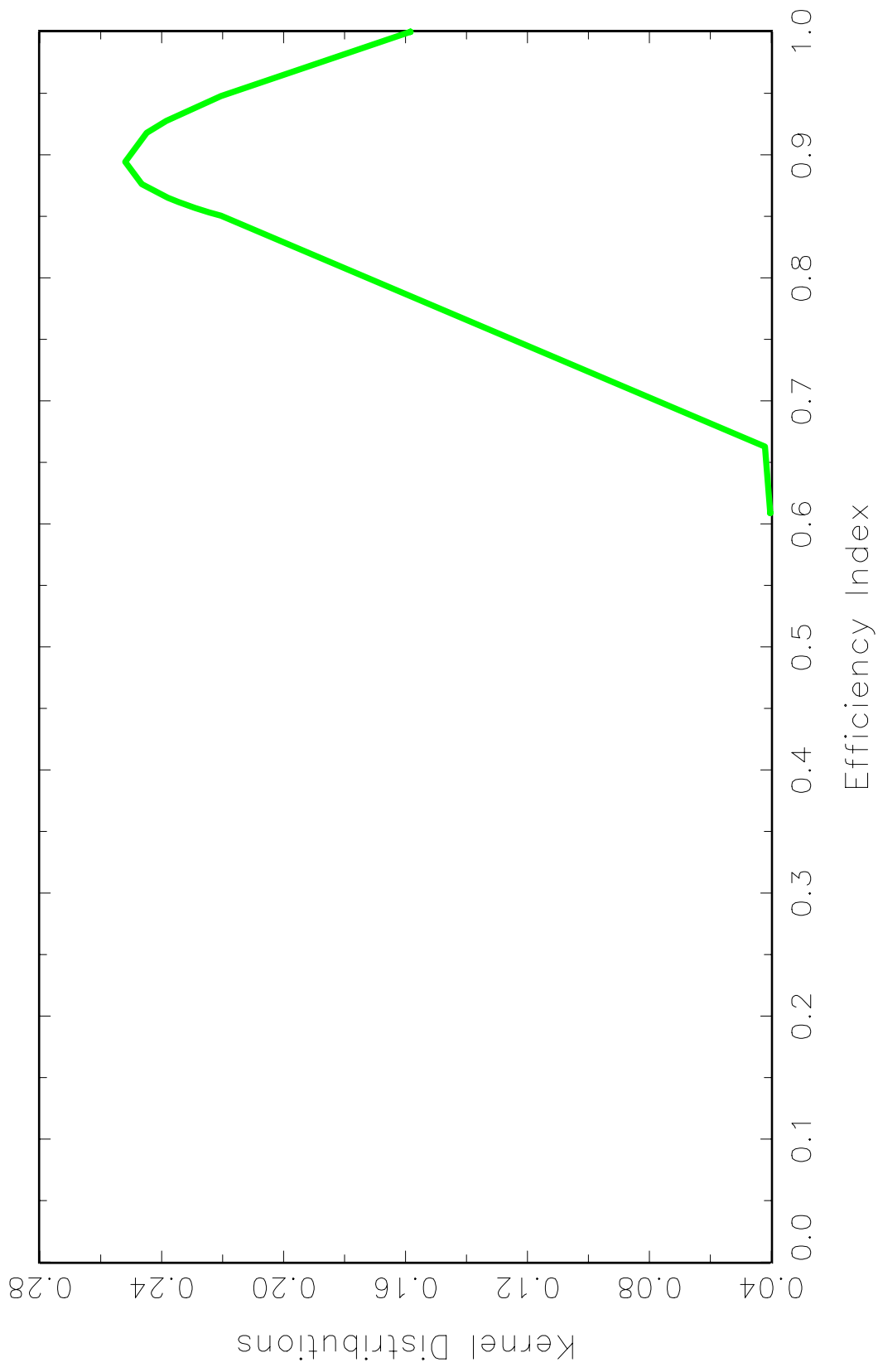Figure 7
1983 Distribution of the DEA-VRS Efficiency Index

Figure 8

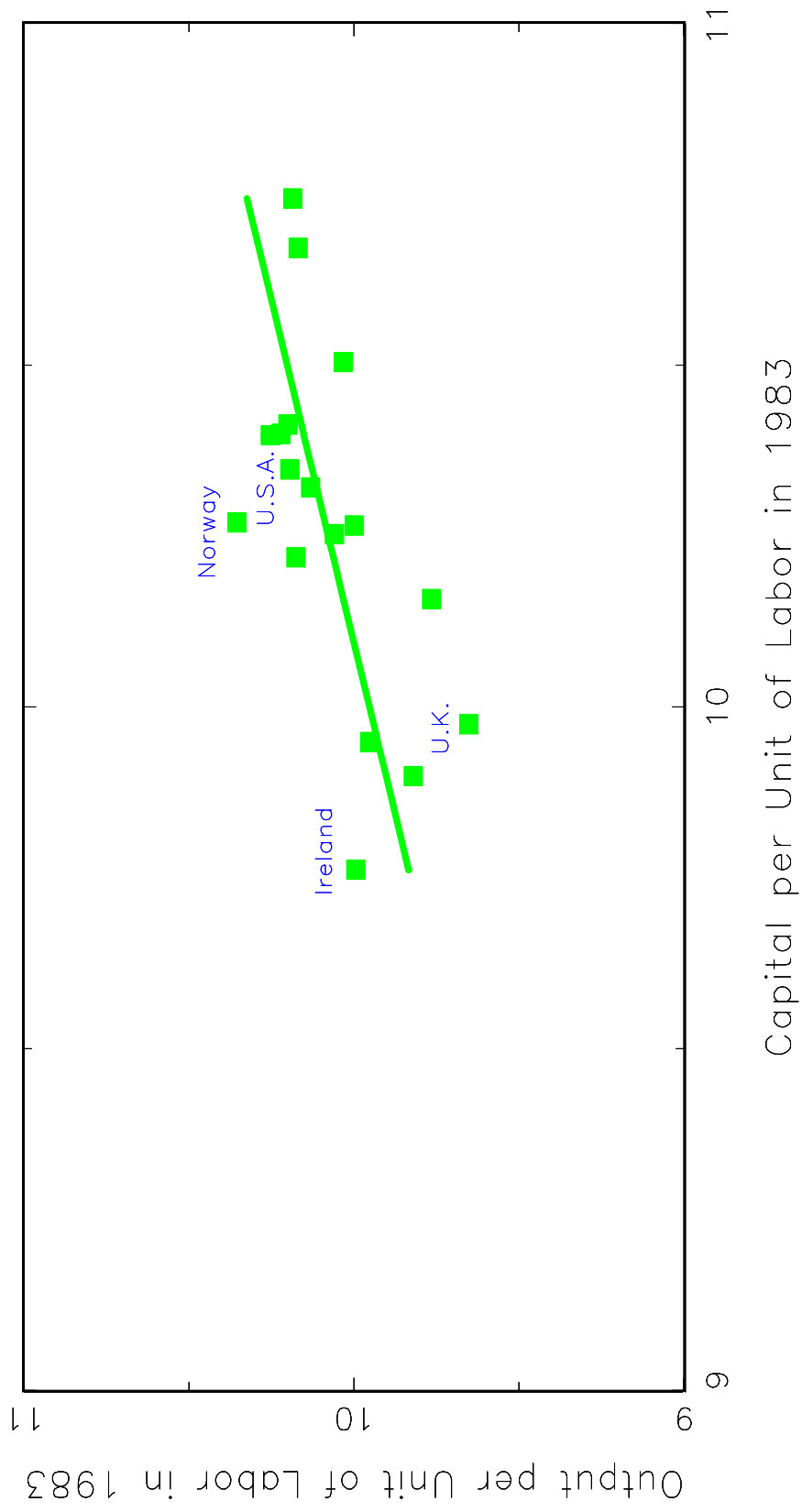1983 Parametric-CRS Production Frontier (In Logs)

Output per Unit of Labor in 1983

Capital per Unit of Labor in 1983

Figure 9

1983 Nonparametric–CRS Production Frontier (In Logs)

Capital per Unit of Labor in 1983

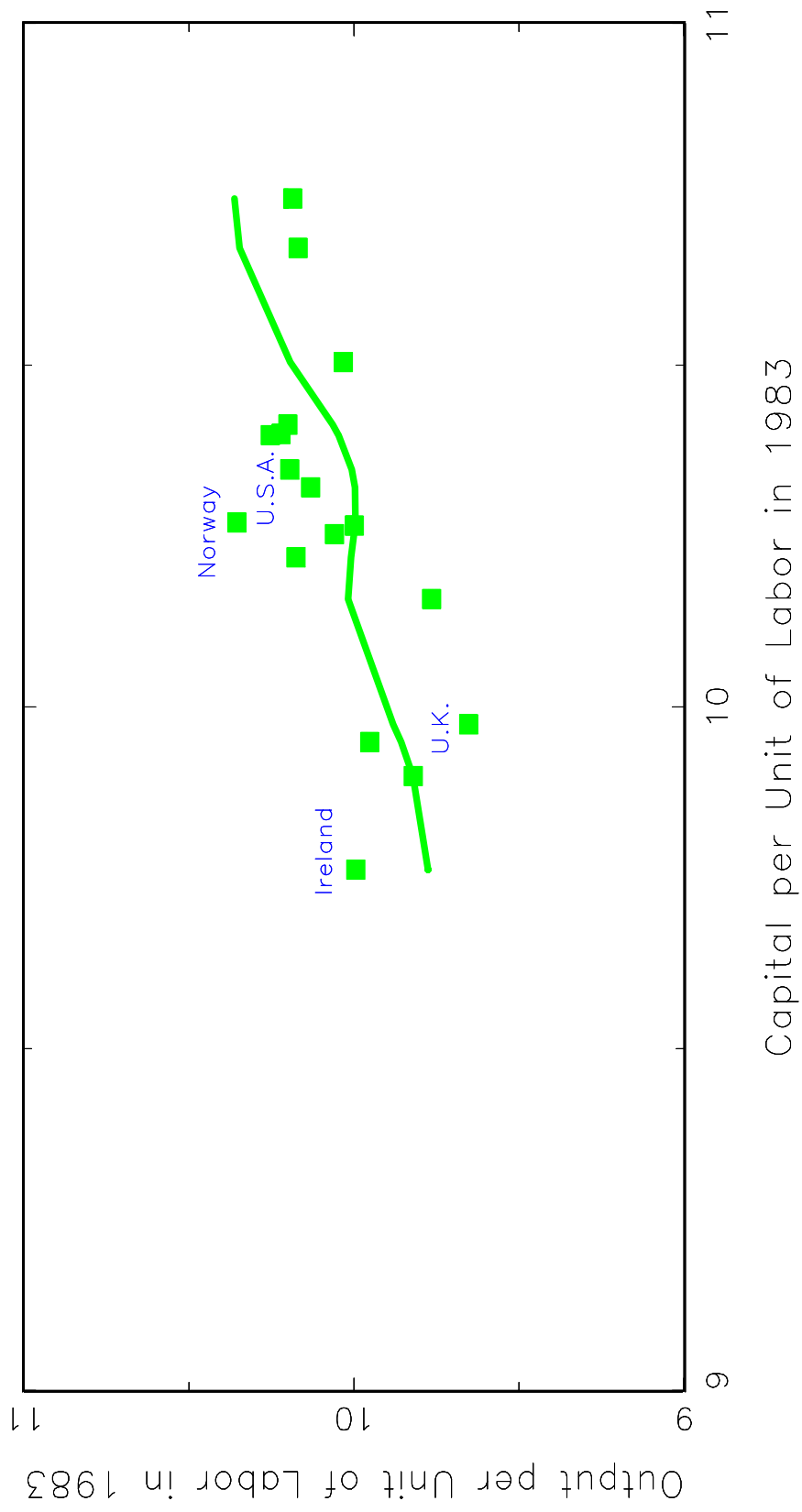Output per Unit of Labor in 1983

Norway

U.S.A.

Ireland

U.K.

Figure 10

1983 DEA-CRS Production Frontier (In Ten-Thousands)

Capital per Unit of Labor in 1983

Output per Unit of Labor in 1983

U.S.A.

U.K.

Ireland