Metafrontier Functions for the Study of Inter-regional Productivity Differences

D.S. PRASADA RAO and CHRISTOPHER J. O'DONNELL
School of Economics, University of Queensland, Brisbane, QLD 4072 Australia

and

GEORGE E. BATTESE
School of Economics, University of New England, Armidale, NSW 2351, Australia

Abstract

This paper uses the concept of the metafrontier function to study regional differences in production technologies. The paper has three components. The first deals with the analytical framework necessary for the definition of a metafrontier function. The second component studies the properties of a metafrontier function estimated using non-parametric data envelopment analysis (DEA). The third component focuses on the estimation of a metafrontier function within the parametric framework of stochastic frontier analysis (SFA). The empirical application of the models uses cross-country agricultural sector data. The metafrontiers that are estimated by DEA and SFA are presented and discussed.

JEL Classifications: C23, C51, C63, D24, L6

1. Introduction

It is a common practice to use production frontiers to assess the level of efficiency of production units or firms. Such frontiers are identified using non-parametric or parametric methods predicated on various non-stochastic and stochastic assumptions. Once a frontier surface is determined, the efficiency of each firm is measured relative to the frontier using radial measures of efficiency that were proposed by Farrell (1957). Frontiers are estimated using cross-sectional or panel data on the levels of inputs used and outputs produced by the firms. While technical efficiencies of firms that are measured with respect to a given frontier are comparable, this is not normally the case among firms that operate under different technologies. Such problems arise when comparisons of firms from different regions are involved. Battese and Rao (2002) and Battese, Rao and O'Donnell (2004) provide frameworks for comparisons when efficiency is measured using stochastic frontier models. While these papers outline a workable approach to providing regional comparisons, they do not examine the analytical framework necessary to undertake such comparisons.

The principal objective of this paper is to establish a framework for metafrontiers based on the axioms associated with production technologies. The metafrontier concept used in this paper is based on the concept of the metaproduction function defined by Hayami and Ruttan (1971, p. 82): "the metaproduction function can be regarded as the envelope of commonly conceived neoclassical production functions." We use both data envelopment analysis (DEA) and stochastic frontier analysis (SFA) to compare efficiencies of firms in different regions that operate under different technologies. DEA is the most popular non-parametric and non-stochastic approach to efficiency measurement. The empirical part of the paper focuses on inter-region comparisons of productivity in agriculture using country-level data drawn from the Food and Agriculture Organization (FAO) of the United Nations. Countries are grouped into regions using the standard geographical classification of countries.

2. Analytical Framework

Efficiency measurement of firms or production units is deeply rooted in the theory of production and the concept of distance functions. In this section, we present the notation and concepts underlying efficiency measurement.

Let y and x, column vectors of dimensions M and N, respectively, denote the nonnegative real numbers for output and input vectors of a firm. We consider the case where there are K (>1) regions, and firms in each region operate under a region-specific technology, T^k (k = 1, 2, …, K).

Since technology is a representation of the state of knowledge pertaining to the transformation of N inputs into M outputs, it is possible to conceptualise the existence of an over-arching technology, referred to as the metatechnology, which we represent by T^*.

The technology in a given region can be defined using either the *technology set* or the *output* or *input sets*. The technology set consists of all output vectors that can be produced using a nonnegative vector of inputs. It is defined as:

$$T = \{ (x, y) : x \text{ and } y \text{ nonnegative}; x \text{ can produce } y \}.$$

The output sets, which are similar to the production possibility sets, are defined for any input vector, x, as:

$$P(x) = \{y : (x, y) \in T\}.$$

The boundary of the output set shows the production possibility frontier.

The input sets are defined for any output vector, y, as:

$$L(y) = \{x : (x, y) \in T\}.$$

The boundaries of the input sets define the “isoquants”.

Following Färe and Primont (1995), the following axioms relating to the production technology are specified using the output sets. For region k technology, T^k, we assume that

(1) $0 \in P^k(x)$ (it is possible to produce nothing);

(2) For all x, $y \in P^k(x)$, and if $0 < \theta \leq 1$, then $y^* = \theta y \in P^k(x)$ (weak disposability);

(3) For all x, $P^k(x)$ is a closed and bounded set; and

(4) For all x, $P^k(x)$ is convex.

In addition, it is common to assume input convexity and weak disposability of inputs. These axioms are necessary for obtaining solutions to revenue maximisation and cost minimisation problems.

The main focus of the paper is to measure the technical efficiency of an observed input and output combination (x, y) relative to the technology of a given region k. Output and input distance functions are used in measuring technical efficiency. These are briefly defined below.

Output Distance Function

Let $D_o^k(x,y)$ represent the output distance function for region k technology, which is defined as:

$$D_o^k(x,y) = \inf_\theta \left\{\theta > 0 : (y/\theta) \in P^k(x)\right\}. \quad (1)$$

From the definition of the distance function, and using the axioms on the technology, it is easy to verify that:

i. $D_o^k(x,y) < 1$ if y belongs to the interior of $P^k(x)$;

ii. $D_o^k(x,y) = 1$ if y belongs to the boundary of $P^k(x)$; and

iii. if y is outside the set $P^k(x)$ then it needs to be scaled down so that it is feasible to produce it using x.

An observation (x, y) can be considered technically efficient (i.e., y belongs to the production frontier of $P^k(x)$) if and only if $D_o^k(x,y)=1$. In addition, the output distance function is homogeneous of degree one in outputs.

Input Distance Function

Let $D_i^k(x,y)$ denote the input distance function defined using the technology of region k, which is defined by:

$$D_i^k(x,y) = \sup_{\lambda}\left\{\lambda > 0 : (x/\lambda) \in L^k(y)\right\}. \qquad (2)$$

For any input vector, x, belonging to the interior of L(y), $D_i^k(x,y) > 1$, and, if x belongs to the boundary, then $D_i^k(x,y) = 1$. The input distance function shows the maximum degree to which a given input vector can be radially contracted and yet produce the same output vector, y. The input distance function is homogeneous of degree one in inputs.

The output and input distance functions are often used to characterise the production technology (see Färe and Primont, 1995).

Metaproduction Technology

We consider technology to be a state of knowledge in existence at a given point of time. What we have described so far in the form of regional technologies is a description of some of the elements of this knowledge. Therefore, we define the metatechnology as the totality of the regional technologies. For example, if a particular output, y, can be produced using a given input vector, x, in any one of the regions, we consider that (x, y) belongs to the metatechnology, T^*, that is defined by:

T^* = {(x, y): x ≥ 0 and y ≥ 0, such that x can produce y in at least one regional technology, $T^1, T^2, \ldots, T^K$}.

From this definition, it follows that $T^* \supseteq \{T^1 \cup T^2 \cup \ldots \cup T^K\}$.

We need to ensure that the metatechnology, T^*, satisfies all the necessary technology axioms. If the metatechnology is defined using a finite number of regions, and, if all the regional technologies satisfy all the production axioms, then it is easy to prove that T^* also satisfies all the production axioms except the convexity axiom.

In order to ensure the convexity property, we define the metatechnology as the convex hull of the union of region-specific technologies, denoted by

$$T^* \equiv \text{Convex Hull } \{T^1 \cup T^2 \cup ... \cup T^K\}.$$

Let $D_o^*(x,y)$ and $D_i^*(x,y)$ denote the output- and input-distance functions defined using the metatechnology, T^*. Following the definition of the metatechnology, we can easily establish the following results:

Result 1: For any given region k, $D_o^k(x,y) \geq D_o^*(x,y)$, k = 1, 2, ..., K; and

Result 2: For any given region k, $D_i^k(x,y) \leq D_i^*(x,y)$, k = 1, 2, ..., K.

These results follow from the fact that the output and input sets for any particular region are subsets of the corresponding sets from the metatechnology.

Technical Efficiency and Technology Gap Ratio

From Results 1 and 2, whenever a strict inequality is observed, we can obtain a measure of the gap between the region k technology and the metatechnology, as outlined below.

Output-orientated Measures

The technical efficiency of an observed pair (x, y) with respect to the technology of region k is defined as:

$$TE_o^k(x,y) = D_0^k(x,y). \tag{3}$$

For example, if $D_o^k(x, y)$ = 0.8 then the technical efficiency measure indicates that the output vector, y, is 80 per cent of the potential output, using the same input vector.

The output-orientated *technology gap ratio* can be defined using the output distance functions from technologies T^k and T^* as:

$$TGR_o^k\ (x, y) = \frac{D_o^*(x, y)}{D_o^k(x, y)}. \tag{4}$$

Using the definition of output-orientated technical efficiency, the technology gap ratio is expressed by:

$$TGR_o^k\ (x, y) = \frac{TE_o^*(x, y)}{TE_o^k(x, y)}. \tag{5}$$

In the example above, if the technical efficiency of (x, y) with respect to the metatechnology is 0.6 then the technology gap ratio is 0.75 (= 0.6/0.8). This means that, given the input vector, the potential output vector for region k technology is 75 per cent of that represented by the metatechnology.

Equation (5) provides a convenient decomposition of the technical efficiency of a particular input-output combination evaluated at the metatechnology, relative to that of region k, namely:

$$TE_o^*(x, y) = TE_o^k(x, y) \times TGR_o^k(x, y) \tag{6}$$

which shows that technical efficiency measured with reference to the metatechnology can be decomposed into the product of the technical efficiency measured with reference to the region k technology (representing the existing state of knowledge) and the technology gap ratio between the region k technology and the metatechnology.

Input-orientated Measures

The input-orientated technical efficiency of an observed pair (x, y) with respect to the technology of region k is defined as:

$$TE_i^k(x,y) = \frac{1}{D_i^k(x,y)}. \tag{7}$$

If the measured input-orientated technical efficiency of a given (x, y) is 0.7 then it implies that y can be produced using 70 per cent of the input vector, x.

The input-orientated *technology gap ratio* can be defined using the output distance functions from technologies T^k and T^* as:

$$TGR_i^k(x,y) = \frac{D_i^k(x,y)}{D_i^*(x,y)} = \frac{TE_i^*(x,y)}{TE_i^k(x,y)}. \tag{8}$$

These ratios are always greater than or equal to zero and less than or equal to one; and equality at one holds when the regional technology frontier coincides with the metatechnology frontier for the input and output vectors, x and y. The input-orientated technology gap ratio can be decomposed along the lines of equation (6).

Empirical measurement of these various measures of technical efficiency and technology gap ratios requires an empirical description of the technology derived from a set of observations on inputs and outputs for a random sample of firms. Once the data are available, one may estimate the frontier using a non-parametric and non-stochastic technique, such as data envelopment analysis (DEA), or a parametric stochastic approach, such as stochastic frontier analysis (SFA). In the following two sections, we describe our methodology based on these two approaches.

3. DEA Approach to Metafrontiers

DEA is a linear-programming methodology that uses data on inputs and outputs of firms, production entities or decision-making units (referred to as *units* in this and the next section) to construct a piece-wise linear surface over the data points. This frontier surface is constructed by the solution of a sequence of linear programming problems – one for each unit in the sample. The degree of technical inefficiency of each unit (the distance between the observed data point and the frontier) is produced as a by-product of the frontier construction method.

In simple terms, DEA attempts to benchmark the performance of each unit against the best practice for all units. The "best-practice" is derived after taking into account the output structure (e.g., share of crops and livestock in total output) as well as the input structure of the unit under consideration. A measure of technical efficiency is then obtained by measuring the radial distance of the unit from its best practice. The technique also identifies the units in the data set that define the best practice – such units are referred to as "peers". It is also possible to measure the importance of each of the peers through a set of weights derived through the application of the DEA technique (see Coelli, Rao and Battese (1998) for details).

DEA can be either input-orientated or output-orientated. In the input-orientated case, DEA defines the frontier by seeking the maximum possible proportional reduction in input usage, with output levels held constant, for each unit. In the output-orientated case, the DEA method seeks the maximum possible proportional increase in outputs, with input levels held fixed. The two measures provide the same technical efficiency scores when a constant returns-to-scale (CRS) technology applies, but are unequal when variable returns-to-scale (VRS) technology holds.

Given that units are within *regions*, it is possible to identify a "regional frontier" using DEA on the data for units from the given region. Thus, DEA can be used to construct K regional frontiers. The metafrontier is then constructed by using DEA to analyse the data set obtained by pooling all the observations for units from all the regions.

Construction of Regional Frontiers

If region k consists of data on L_k units, the linear programming (LP) problem that is solved for the i-th unit in an output-orientated DEA model is as follows:

$$\max_{\phi,\lambda} \ \phi,$$

$$\text{such that} \quad -\phi y_i + Y_k\lambda \geq 0,$$
$$x_i - X_k\lambda \geq 0,$$
$$\lambda \geq 0, \qquad (9)$$

where

y_i is the M×1 vector of output quantities for the i-th unit;

x_i is the N×1 vector of input quantities for the i-th unit;

Y_k is the M×L_k matrix of output quantities for all L_k units;

X_k is the N×L_k matrix of input quantities for all L_k units;

λ is a L_k×1 vector of weights; and

ϕ is a scalar.

Observe that ϕ will take a value greater than or equal to one, and that ϕ-1 is the proportional increase in outputs that could be achieved by the i-th unit, with input quantities held constant. Note also that 1/ϕ defines a technical efficiency (TE) score which varies between zero and one (this is the output-orientated TE score reported in our empirical application in Section 6).

The above LP is solved L_k times; once for each unit in region k. Each LP produces ϕ and λ vectors. The ϕ-vector provides information on the technical efficiency score for the i-th unit, and the λ-vector provides information on the *peers* of the (inefficient) i-th unit. The peers of the i-th unit are those efficient units in the region that define the facet of the frontier against which the (inefficient) i-th unit is projected.

Construction of the Metafrontier

The metafrontier is constructed using a DEA model based on the pooled data for all the units in all the regions. Since we have a total of $L = \sum_k L_k$ units, we re-run the above LP model with the input and output matrices with data for all units:

$$\max_{\phi^*,\lambda^*} \phi^*,$$

$$\text{such that} \quad -\phi^* y_i + Y^*\lambda^* \geq 0,$$

$$x_i - X^*\lambda^* \geq 0,$$

$$\lambda^* \geq 0, \qquad (10)$$

where

y_i is the M×1 vector of output quantities for the i-th unit;

x_i is the N×1 vector of input quantities for the i-th unit;

Y^* is the M×L matrix of output quantities for all the L units;

X^* is the N×L matrix of input quantities for all the L units;

λ^* is the L×1 vector of weights; and

ϕ^* is a scalar.

The optimum solution of (10) provides a technical efficiency score for a given unit relative to the metafrontier identified using data from all the units in all regions.

It is easy to see that, for any given unit, ϕ^* is no larger than ϕ, because the constraints in the regional LP problem (9) are a subset of the constraints in the metafrontier LP problem (10). Units are shown to be not more technically efficient when they are assessed against the metafrontier than against the regional frontier.

4. A Stochastic Metafrontier Model

Suppose that the inputs and outputs for units in a given industry (the single output case is considered in this section) are such that stochastic frontier production function models are appropriate for the K different regions. Suppose that, for the k-th region, there are sample data on L_k units that produce the one output involved using the various inputs, and that the stochastic frontier model for this region is defined by

$$Y_{it(k)} = f\left(x_{it(k)}, \beta_{(k)}\right)e^{V_{it(k)}-U_{it(k)}}, \; i = 1, 2, \ldots, L_k; \; t = 1, 2, \ldots, T; \; k=1, 2, \ldots, K, \qquad (10)$$

where $Y_{it(k)}$ denotes the output for the *i*-th unit in the *t*-th time period for the k-th region; $x_{it(k)}$ denotes a vector of values of functions of the inputs used by the *i*-th unit in the *t*-th time period for the k-th region; $\beta_{(k)}$ denotes the parameter vector associated with the x-variables for the stochastic frontier for the k-th region; the $V_{it(k)}$s are assumed to be identically and independently distributed as $N\left(0, \sigma^2_{v(k)}\right)$-random variables, independent of the $U_{it(k)}$s, which are defined by the truncation (at zero) of the $N\left(\mu_{it(k)}, \sigma^2_{(k)}\right)$-distributions, where the $\mu_{it(k)}$s are defined by some appropriate inefficiency model, e.g., one of the Battese and Coelli (1992, 1995) models. For simplicity of presentation, the model for the k-th region is assumed to be given by

$$Y_{it} = f\left(x_{it}, \beta_{(k)}\right)e^{V_{it(k)}-U_{it(k)}} \equiv e^{x_{it}\beta_{(k)}+V_{it(k)}-U_{it(k)}}. \qquad (11)$$

This expression assumes that the exponent of the frontier production function is linear in the parameter vector, $\beta_{(k)}$, so that x_{it} is a vector of functions (e.g., logarithms) of the inputs for the i-th unit in the t-th time period involved.

The metafrontier production function model for units in the industry is expressed by

$$Y^*_{it} \equiv f(x_{it}, \beta^*) = e^{x_{it}\beta^*}, \; i = 1, 2, \ldots, L = \sum_{k=1}^{K} L_k; \; t = 1, 2, \ldots, T, \qquad (12)$$

where β^* denotes the vector of parameters for the metafrontier function such that

$$x_{it}\beta^{*} \geq x_{it}\beta_{(k)}. \tag{13}$$

The metafrontier function, defined by equations (12)-(13), is a production function of specified functional form that does not fall below the deterministic functions for the stochastic frontier models of the regions involved. Battese and Rao (2002) give a more extensive literature review and proposed a stochastic metafrontier model that assumes a different data-generation mechanism for the metafrontier than for the different regional frontiers. The model in this paper assumes that data-generation models are only defined for the stochastic frontier models for the units in the different regions. The metafrontier is assumed to be a smooth function and not a segmented envelope of the stochastic frontier functions for the different regions. Figure 1 is the graph of the metafrontier function from Battese, Rao and O'Donnell (2004).

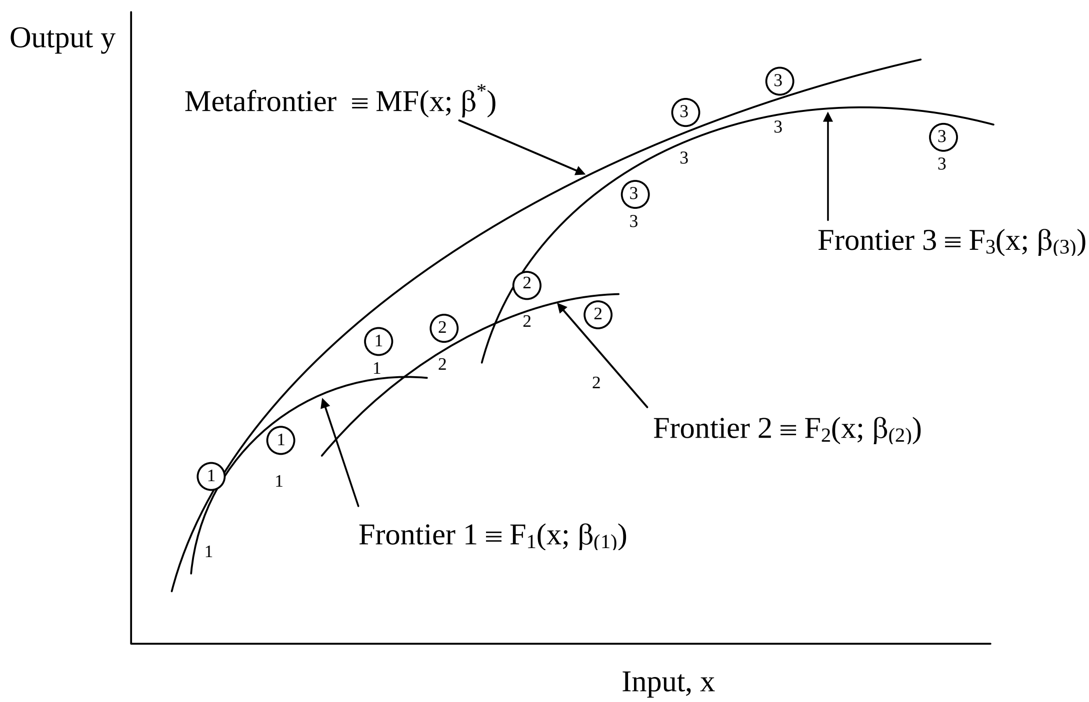

Figure 1: Metafrontier Function Model

The observed output for the i-th unit at the t-th time period, defined by the stochastic frontier for the k-th region in equation (11), is alternatively expressed in terms of the metafrontier function of equation (12) by

$$Y_{it} = e^{-U_{it(k)}} \times \frac{e^{x_{it}\beta_{(k)}}}{e^{x_{it}\beta^{*}}} \times e^{x_{it}\beta^{*}+V_{it(k)}} \tag{14}$$

where the first term on the right-hand side of equation (14) is the technical efficiency relative to the stochastic frontier for the k-th region,

$$TE_{it}^{k} = \frac{Y_{it}}{e^{x_{it}\beta_{(k)}+V_{it(k)}}} = e^{-U_{it(k)}} . \tag{15}$$

The second term on the right-hand side of equation (14) is the technology gap ratio (TGR) for the i-th unit (in the k-th region) at the t-th time period:

$$TGR_{it}^{k} = \frac{e^{x_{it}\beta_{(k)}}}{e^{x_{it}\beta^{*}}} . \tag{16}$$

This measures the ratio of the output for the frontier production function for the k-th region relative to the potential output that is defined by the metafrontier function, given the observed inputs. The technology gap ratio has values between zero and one because of equation (13).

The *technical efficiency of the i-th unit*, given the t-th observation, *relative to the metafrontier*, denoted by TE_{it}^{*}, is defined in an analogous way to equation (15). It is the ratio of the observed output relative to the last term on the right-hand side of equation (14), which is the metafrontier output, adjusted for the corresponding random error, i.e.,

$$TE_{it}^{*} = \frac{Y_{it}}{e^{x_{it}\beta^{*}+V_{it(k)}}} . \tag{17}$$

Equations (14)-(17) imply that an alternative expression for the technical efficiency relative to the metafrontier is given by

$$TE_{it}^{*} = TE_{it}^{k} \times TGR_{it}^{k} . \tag{18}$$

Thus the technical efficiency relative to the metafrontier function is the product of the technical efficiency relative to the stochastic frontier for the given region and the technology gap ratio for that region. Because both the latter measures are between zero and one, the technical efficiency relative to the metafrontier is also between zero and one, but is less than the technical efficiency relative to the stochastic frontier for the region of the unit.

The parameters and measures associated with the metafrontier model of equations (11)-(13) can be estimated as follows:

1) Obtain the maximum-likelihood estimates, $\hat{\beta}_{(k)}$, for the $\beta_{(k)}$-parameters of the stochastic frontier for the *k*-th region using, for example, the FRONTIER program (Coelli, 1996a);
2) Obtain estimates, $\hat{\beta}^*$, for the β^*-parameters of the metafrontier function such that the estimated function best envelopes the deterministic components of the estimated stochastic frontiers for the different regions. (A formal procedure is defined below.)
3) Estimates for the technical efficiencies of units relative to the metafrontier function can be predicted by

$$T\hat{E}_{it}^{*} = T\hat{E}_{it}^{k} \times TG\hat{R}_{it}^{k} \qquad (19)$$

where $T\hat{E}_{it}^{k}$ is the predictor for the technical efficiency relative to the k-th regional frontier, as proposed in Battese and Coelli (1988, 1992), which is programmed to be calculated in FRONTIER; and $TG\hat{R}_{it}^{k} = \frac{e^{x_{it}\hat{\beta}_{(k)}}}{e^{x_{it}\hat{\beta}^*}}$ is the estimate for the technology gap ratio for the i-th unit in the k-th region relative to the industry potential, obtained by using the estimates for the parameters involved.

In our empirical application, the β^*-parameters of the metafrontier function are estimated by solving the optimisation problem below:

$$\text{Minimise} \quad \sum_{t=1}^{T} \sum_{i=1}^{L} \left| (ln\, f(x_{it}, \beta^*) - ln\, f(x_{it}, \hat{\beta}_{(k)}) \right| \qquad (20)$$

subject to:

$$ln\, f(x_{it}, \beta^*) \geq ln\, f(x_{it}, \hat{\beta}_{(k)}). \qquad (21)$$

where $ln\, f(x_{it}, \hat{\beta}_{(k)})$ is the logarithm of the estimated deterministic component of the stochastic frontier for the k-th region, associated with the translog production function that is used in the empirical application.

Given that $f(x_{it}, \beta^*)$ in equation (12) is log-linear in the parameters (as assumed in this paper), the optimisation problem in (20)-(21) is solved by the linear programming (LP) problem:

Minimise $$\sum_{i=1}^{L} \sum_{t=1}^{T} (x_{it}\beta^* - x_{it}\hat{\beta}_{(k)}) \qquad (22)$$

subject to:

$$x_{it}\beta^* \geq x_{it}\hat{\beta}_{(k)}, \qquad (23)$$

for all k = 1, 2, …, K.

The solution to the above problem is equivalently obtained by minimising the objective function, $\bar{x}\beta^*$, subject to the linear restrictions of equation (23), where $\bar{x}$ is the row vector of means of the elements of the x_{it}-vectors for all units in the data set. This follows because the estimates of the stochastic frontiers for the different regions, $\hat{\beta}_{(k)}$, k = 1, 2,..., K, are assumed to be fixed for the linear programming problem.

The above formulation and the one involving minimisation of the sum of squared deviations are discussed in more detail in Battese, Rao and O'Donnell (2004). Standard errors for the estimators for the metafrontier parameters can be obtained using simulation or bootstrapping methods.

5. Data for Inter-regional Productivity Comparisons

The present study is based on data exclusively drawn from the FAOSTAT system of statistics used for dissemination of statistics compiled at the Statistics Division of the Food and Agriculture Organization in Rome. The Statistics Division maintains a fairly regularly updated website where data on agriculture are made available to the potential users. The site can be reached through the general FAO URL: **www.fao.org**

The study covers 97 countries that are major agricultural producers of the world. These countries account for roughly 99 per cent of the world's agricultural output as well as 99 per cent of the world's population. The countries included in the data set are evenly distributed over all the regions of the world. The 97 countries in the data set are grouped into four regions (see Table 1):

Region 1	Africa	27 countries
Region 2:	The Americas	21 countries
Region 3	Asia	26 countries
Region 4	Europe	23 countries

We used data from 1986 to 1990. A total of $97 \times 5 = 485$ observations were available to estimate various DEA frontiers. Only 483 observations were available to estimate the stochastic frontiers – two observations were deleted because the fertiliser input was zero (and the functional form involved logarithms).[1] The DEA and SFA frontiers, discussed in previous sections, are defined for data on the different countries that are listed in Table 1 (i.e., *units* in earlier sections are considered to be *countries* in the application below).

Output Series (Y): This paper considers two output aggregates, viz., crops and livestock outputs. This distinction is made in order to explicitly recognise the differences in production techniques involved in producing crops in comparison with livestock. While it is true that different agricultural commodities within these broadly defined groups also exhibit differences in their production, it is not possible to treat very disaggregated output data for cross-country productivity comparisons and analysis. Thus, a total of 185 agricultural commodities are broadly classified into crops and livestock products and aggregate series for these two output categories are derived using the steps outlined below. The output aggregates used here refer exclusively to the final output (agricultural output net of feed and seed) in different countries for these two commodity groups.

[1] The approach suggested by Battese (1997) is not used to include these zero observations.

For the year 1990, the final output aggregates for crops and livestock are from Table 4 in Rao (1993), arising from the 1993 FAO study. These aggregates are constructed using international average prices (expressed in US dollars), derived using the Geary-Khamis method (see Chapter 4 of Rao (1993) for a detailed description of this method) for the benchmark year 1990.[2] The output series for 1990, presented in Rao (1993), are adjusted for price differences across countries, expressed in billions of US dollars.

Input Series*:* Since the application of DEA requires that the numbers of input and output variables be kept at reasonable levels, we consider five important input variables.[3] Detailed descriptions of these variables are given below.

Land (X_1): This variable includes the arable land, land under permanent crops as well as the area under permanent pasture, expressed in millions of hectares. Arable land includes land under temporary crops (double-cropped areas are counted only once), temporary meadows for mowing or pasture, land under market or kitchen gardens and land temporarily fallowed (for less than five years). Land under permanent crops is the land cultivated with crops that occupy the land for long periods and need not be replanted after each harvest. This category includes land under flowering shrubs, fruit trees, nut trees and vines but excludes land under trees grown for wood or timber. Land under permanent pasture is land that is used permanently (for at least five years) for forage crops, either cultivated or growing wild.

Machinery (X_2): This variable includes the total number of wheeled and crawler tractors used in agriculture, but excludes garden tractors. Only the number of tractors is used, with no allowance made for the horsepower of the tractors.

Labour (X_3): The labour variable used is the economically active population in agriculture, which is defined as all persons engaged in or seeking employment in the

[2] The Geary-Khamis international average prices are based on prices (in national currency units) and quantities of 185 agricultural commodities in 103 countries.
[3] Choice of the number of outputs and inputs used in the assessment of total factor productivity essentially relates to the availability of degrees of freedom (see Coelli, Rao and Battese (1998) for details).

operation of a family farm or business, whether as employers, own-account workers, salaried employees or unpaid workers. The economically active population in agriculture includes all persons engaged in economic activities in agriculture, forestry, hunting or fishing. This variable obviously overstates the labour input used in agricultural production, the extent of overstatement depends upon the level of development of the country.[4] Since we examine the agricultural sector productivity for a country, as a whole, it is quite appropriate that the economically active population is used as an aggregate measure of the labour input into the sector. Further refinements would require information on differentials in skill levels and the numbers of hours worked on the farms to be available.

Fertiliser (X_4): This input is quite difficult to measure. The FAO Fertiliser Yearbook provides details of fertiliser production and use in different countries, and the data available involve a large number of fertilisers. It is impossible to consider fertiliser data in such detail. Thus, following other studies (Hayami and Ruttan, 1970; Fulginiti and Perrin, 1997) on inter-country comparisons of agricultural productivity, we use the sum of the nitrogen (N), potassium (P_2O_2) and phosphate (K_2O) contained in the commercial fertilisers that were applied as the measure of fertiliser input in this paper. This variable is expressed in thousands of tonnes.

Livestock (X_5): The livestock input variable used in the study is the sheep-equivalent of five categories of animals. The categories of animals considered are buffaloes, cattle, pigs, sheep and goats. Data on numbers of these animals are converted into sheep equivalents using the following conversion factors: 8.0 for buffalo and cattle; and 1.0 for sheep, goats and pigs.[5] Due to their short life span, the numbers of chickens are not included in the livestock figures.

A complete list of the countries, classified into the four regions, is provided in Table 1. Australia, New Zealand and Papua New Guinea are included in the Asian region. The former USSR is included in Europe. Table 2 provides descriptive statistics for all the variables by region. It is evident from these statistics that these regions consist of countries differing in size and also by the land and labour endowments.

[4] There could be a significant percentage of labour (as defined here) in disguised unemployment.
[5] These conversion figures correspond very closely with those used in Hayami and Ruttan (1970).

6. Empirical Results

Estimates of the parameters of the DEA and SFA models that are described in previous sections are presented and discussed. The acronyms for the models are defined as follows:

DEA-REG	DEA estimates of the regional frontiers
DEA-MF	DEA estimates of the metafrontier for all regional data
SFA-REG	ML estimates of the regional stochastic frontiers
SFA-MF	LP estimate of the SFA metafrontier
SFA-POOL	ML estimate of the stochastic frontier for all regional data.

The DEA estimates were obtained using DEAP 2.1 (see Coelli, 1996b). DEA results were obtained using CRS, a single-stage estimation method and an input-orientation. (DEAP had trouble estimating the pooled model when an output orientation was used or when a two-stage estimation technique was used.) Under the assumption of constant returns to scale it does not make much difference between input- and output-orientated models.

The SFA estimates were obtained using the stochastic frontier model of Battese and Coelli (1992) and a translog functional form:

$$\ln(y_{it}) = \beta_0 + \sum_{j=1}^{M}\beta_j \ln(x_{jit}) + 0.5\sum_{j=1}^{M}\sum_{k=1}^{M} \beta_{jk}\ln(x_{jit})\ln(x_{kit}) + V_{it} - U_{it} \qquad (24)$$

where

$$U_{it} = \{\exp[-\eta(t-T)]\}U_i \qquad (25)$$

and U_i is defined by the nonnegative truncation of the $N(\mu, \sigma^2)$-distribution.

The SFA-REG and SFA-POOL estimates were obtained using FRONTIER 4.1c (see Coelli, 1996a). The SFA-POOL estimates were used to obtain a test of the null hypothesis that the different regions had the same stochastic frontier models. The SFA-MF estimates were obtained using SHAZAM – the code is presented in the Appendix.

In order to be able to run the DEA models with one output and five inputs and separately for each of the regions, we had to pool the data for each of the regions for the five years, 1986-1990. However, observations for each country for different years are treated as separate observations.

Technical efficiency estimates and estimates of the technology gap ratios (TGR) are summarised in Table 3. Results for selected countries are reported in Table 4. SFA parameter estimates are reported in Table 5.

Efficiency using DEA Frontiers

The DEA analysis produces some interesting output. The technique identifies the best performing countries, known as peers, for each of the regions. Since we consider each country in a given year as a different observation, it is possible to observe a country say the Netherlands in 1986 as well as 1990 as two peers for the DEA frontier for Europe.

Peer or best-performing countries in different regions are listed in Table 6. An interesting feature here is that a majority of the peer countries refer to 1989 and 1990. If agricultural technology shows improvement over time, one would expect peers from the later years to appear on the frontier. In Asia, New Zealand (1986, 1988 and 1989) appeared as a peer. However, the Philippines, Israel and New Zealand together seem to define best practice for many other countries. There are some countries in Asia that are on the frontier but only appear in defining best practice for one country in one year.

Table 3 also provides average DEA scores from the regional and metafrontiers. In Africa, the average technical efficiency is 0.679 in 1990, which is much higher than the corresponding average for SFA (0.504). The technology gap ratio is fairly constant over the five-year period (about 0.89). The higher averages may, in part, be attributable to the fact that several countries (in several years) appear as peers and, therefore, have technical efficiency scores of 1.0.

Differences in Regional Stochastic Frontiers

If the stochastic frontiers across regions do not differ, then it is possible to just use the pooled stochastic frontier. Table 5 presents estimates of the parameters of the stochastic frontiers for all the regions separately and also using pooled data. The generalised likelihood-ratio test statistic for the null hypothesis that the regional frontiers are identical is LR = 229.76. This has *p*-value of 0.000 (using a chi-square distribution with 72 degrees of freedom), so we reject the null hypothesis that the regional frontiers are the same. Based on this, the parameters of the metafrontier are estimated by solving the LP problem discussed in Section 4.

Efficiency using SFA Frontiers

Table 3 provides average technical efficiency scores for each of the regions for the years, 1986-90. For the African region, the average technical efficiency score is about 0.5, indicating agricultural output by about 50 per cent of the potential, given its regional technology. As a result of the time-varying nature of technical inefficiencies, the average technical efficiency scores show marginal increases over time. But the more interesting feature is the difference between the average technical efficiency scores from the regional and metafrontier models. For example, in 1990, the average SFA technical efficiency of the African region relative to the metatechnology is only 0.366. This suggests that countries in the African region are much less efficient relative to the metafrontier. In fact, even if all the countries in Africa achieved best practice with respect to the technology observed in Africa, the African technology lags behind global technology with a technology gap ratio of 0.773 in 1990.

In the Americas region, the mean efficiency is quite high with respect to its own regional frontier (0.810 in 1990) but is only 0.598 when assessed against the global technology. A similar picture emerges for Asia and Europe.

Efficiency Estimates for Selected Countries

Table 4 provides estimates of technical efficiency for selected countries from different regions. These were picked to see if the results support prior expectations. As expected, Australia is highly efficient with a score around 0.95 under both SFA and DEA models. Our prior expectation was that Australia would be a peer, suggesting it would be a best-performing country in Asia. However, our results reveal that Australia is just below the frontier and that New Zealand and Israel define best practice for Australia. A technology gap ratio of 1.0 suggests that the best practice for that country is also on the global frontier. Results for the United States, the United Kingdom, Canada and the Netherlands provide no surprises.

However, results for both India and China are surprising. This is for two reasons. Both have very low DEA technical efficiency scores that are about 0.35, suggesting the possibility of vast improvements in agriculture. In the 1980s and 1990s, agriculture in these countries, especially in India, vastly improved. The low DEA scores were attributed to the fact that the land and labour inputs may have been overstated in the FAO statistics. Labour is measured using the population that is economically active in agriculture. There is scope for disguised unemployment. The more surprising aspect of this result is the divergence between the efficiency scores from the SFA models. These averages are around 0.93, suggesting the possibility for only modest improvements. But the high scores for these two countries, even under the global technology, need further examination.

7. Conclusions

This paper has developed the concept of metafrontiers for the purpose of studying differences in productivity across different regions and groupings. The metafrontier represents the overall state of knowledge that is only partially revealed by the frontiers from different regions. Since efficiency of a country in a region (or a firm in a sub-group) is assessed against its own frontier, it may not provide an accurate measure of the potential gains through improvements in efficiency.

This paper develops the metafrontier concept using alternative descriptions of production technologies. In the empirical application, we used DEA and SFA to estimate metafrontiers for countries in different regions. Both of these methods are popular in efficiency measurement literature. DEA can be used in studying multi-output and multi-input technologies. DEA treats all noise as inefficiency. It is therefore sensitive to outliers. The SFA results are more satisfactory from this angle but SFA can only handle a single output, thus requiring a certain level of aggregation on the output side. In the context of agriculture, a two-output model using crops and livestock would have been more realistic.

Empirical results derived using cross-country agricultural sector data for the five years, 1986-1990, yielded fairly meaningful results. There are a few specific cases that require further scrutiny.

References

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, 153-169.

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

Battese, G.E. and D.S.P. Rao (2002) "Technology Gap, Efficiency and a Stochastic Metafrontier Function", *International Journal of Business and Economics,* 1 (2), 1-7.

Battese, G.E., D.S.P. Rao and C.J. O'Donnell (2004) "A Metafrontier Production Function for Estimation of Technical Efficiencies and Technology Gaps for Firms Operating Under Different Technologies", *Journal of Productivity Analysis,* 21, 91-103.

Coelli, T.J. (1996a) "A Guide to FRONTIER Version 4.1: A Computer Program for Stochastic Frontier Production and Cost Function Estimation", *CEPA Working Papers*, No. 7/96, School of Economics, University of New England, Armidale.

Coelli, T.J. (1996b) "A Guide to DEAP Version 2.1: A Data Envelopment Analysis (Computer) Program", *CEPA Working Papers No. 8/96*, Department of Econometrics, University of New England, Armidale.

Coelli, T., D.S.P. Rao and G.E. Battese (1998) *Introduction to Efficiency and Productivity Analysis*, Kluwer Academic Publishers, Boston.

Färe, R. and D. Primont (1995) *Multi-output Production and Duality: Theory and Applications*, Kluwer Academic Publishers, Boston.

Farrell, M.J. (1957) "The Measurement of Productive Efficiency", *Journal of the Royal Statistical Society*, Series A, General, 120, Part 3, 253-281.

Fulginiti, L.E. and R.K. Perrin (1977) "Declining Productivity in LDC Agriculture", mimeograph, Department of Agricultural Economics, University of Nebraska, Lincoln.

Hayami, Y. and V.W. Ruttan (1970) "Agricultural Productivity Differences Among Countries", *American Economic Review*, 40, 895-911.

Hayami, Y. and V.W. Ruttan (1971) *Agricultural Development: An International Perspective*, Johns Hopkins University Press, Baltimore.

Rao, D.S.P. (1993) *Inter-country Comparisons of Agricultural Output and Productivity*, Social and Economic Development Paper, No. 112, Food and Agriculture Organisation, Rome.

Appendix: SHAZAM program for the Estimation of Metafrontier Parameters and Technology Gap Ratios

```
* parm.txt contains estimated parameters of regional frontiers reported in Table 5
* sfa1.txt contains 135 data observations for region 1
* sfa2.txt contains 105 data observations for region 2
* sfa3.txt contains 128 data observations for region 3
* sfa4.txt contains 115 data observations for region 4

* READ ESTIMATED PARAMETERS OF REGIONAL FRONTIERS
  read (parm.txt) parm / rows = 21 cols = 4
  do # = 1,4
     dim b# 21
     copy parm b# / fcols=#;# tcols = 1;1
  endo

* READ DATA (and construct matrices and vectors)
  smpl 1 483
  read (sfa1.txt) region t ly lx1-lx5 lx11-lx15 lx22-lx25 lx33-lx35 lx44-lx45 lx55
  smpl 136 240
  read (sfa2.txt) region t ly lx1-lx5 lx11-lx15 lx22-lx25 lx33-lx35 lx44-lx45 lx55
  smpl 241 368
  read (sfa3.txt) region t ly lx1-lx5 lx11-lx15 lx22-lx25 lx33-lx35 lx44-lx45 lx55
  smpl 369 483
  read (sfa4.txt) region t ly lx1-lx5 lx11-lx15 lx22-lx25 lx33-lx35 lx44-lx45 lx55
  smpl 1 483
  genr one = 1
  matrix x = one|lx1|lx2|lx3|lx4|lx5|lx11|lx12|lx13|lx14|lx15|lx22|lx23|lx24|lx25| &
             lx33|lx34|lx35|lx44|lx45|lx55
  stat x / means = xbar
  dim x1 135 21 x2 105 21 x3 128 21 x4 115 21
  copy x x1 / frows=1;135 trows=1;135
  copy x x2 / frows=136;240 trows=1;105
  copy x x3 / frows=241;368 trows=1;128
  copy x x4 / frows=369;483 trows=1;115
  do # = 1,4
     matrix yhat# = x#*b#
  endo

* OBTAIN AND PRINT PARAMETERS OF THE METAFRONTIER
  matrix c = ((-xbar')|xbar')'
  matrix A = (-x)|x
  matrix b = -(yhat1'|yhat2'|yhat3'|yhat4')'
  lp c A b / iter = 5000 primal = bstar
  dim beta1 21 beta2 21
  copy bstar beta1 / frows=1;21 trows=1;21
  copy bstar beta2 / frows=22;42 trows=1;21
  matrix beta = beta1-beta2
  print beta

* OBTAIN AND PRINT TECHNOLOGY GAP RATIOS
  do # = 1,4
     matrix xbeta# = x#*beta
     matrix tgr# = exp(yhat#)/exp(xbeta#)
     stat tgr#
  endo
  matrix tgr = (tgr1'|tgr2'|tgr3'|tgr4')'
  print tgr
  stop
```

Table 1: Countries and Regions

Country Code	Region Code [a]	Country	Country Code	Region Code [a]	Country
1	1	ALGERIA	51	3	SRI LANKA
2	1	ANGOLA	52	3	CHINA
3	1	BURUNDI	53	3	INDIA
4	1	CAMEROON	54	3	INDONESIA
5	1	CHAD	55	3	IRAN
6	1	EGYPT	56	3	IRAQ
7	1	ETHIOPIA PDR	57	3	ISRAEL
8	1	GHANA	58	3	JAPAN
9	1	GUINEA	59	3	CAMBODIA
10	1	COTE DIVOIRE	60	3	KOREA REP
11	1	KENYA	61	3	LAOS
12	1	MADAGASCAR	62	3	MALAYSIA
13	1	MALAWI	63	3	MONGOLIA
14	1	MALI	64	3	NEPAL
15	1	MOROCCO	65	3	PAKISTAN
16	1	MOZAMBIQUE	66	3	PHILIPPINES
17	1	NIGER	67	3	SAUDI ARABIA
18	1	NIGERIA	68	3	SYRIA
19	1	RWANDA	69	3	THAILAND
20	1	SENEGAL	70	3	TURKEY
21	1	SOUTH AFRICA	71	3	VIET NAM
22	1	SUDAN	72	4	AUSTRIA
23	1	TANZANIA	73	4	BEL-LUX
24	1	TUNISIA	74	4	BULGARIA
25	1	UGANDA	75	4	CZECHOSLOVAK
26	1	BURKINA FASO	76	4	DENMARK
27	1	ZIMBABWE	77	4	FINLAND
28	2	CANADA	78	4	FRANCE
29	2	COSTA RICA	79	4	GERMANY
30	2	CUBA	80	4	GREECE
31	2	DOMINICAN RP	81	4	HUNGARY
32	2	EL SALVADOR	82	4	IRELAND
33	2	GUATEMALA	83	4	ITALY
34	2	HAITI	84	4	NETHERLANDS
35	2	HONDURAS	85	4	NORWAY
36	2	MEXICO	86	4	POLAND
37	2	NICARAGUA	87	4	PORTUGAL
38	2	USA	88	4	ROMANIA
39	2	ARGENTINA	89	4	SPAIN
40	2	BOLIVIA	90	4	SWEDEN
41	2	BRAZIL	91	4	SWITZERLAND
42	2	CHILE	92	4	UK
43	2	COLOMBIA	93	4	YUGOSLAV SFR
44	2	ECUADOR	94	3	AUSTRALIA
45	2	PARAGUAY	95	3	NEW ZEALAND
46	2	PERU	96	3	PAPUA N GUIN
47	2	URUGUAY	97	4	USSR
48	2	VENEZUELA			
49	3	BANGLADESH			
50	3	MYANMAR			

[a] Region codes are: 1 = Africa; 2 = Americas; 3 = Asia; 4 = Europe

Table 2: Summary Statistics on Variables in the Empirical Application

		Numbers of Observations	Mean	Standard Deviation	Minimum	Maximum
Region 1 – Africa						
Y	Output	135	2.290	2.173	0.495	10.607
X1	Land	135	30.263	29.305	1.830	122.900
X2	Machinery	135	16.890	33.611	0.085	169.000
X3	Labour	135	5.083	4.180	0.807	19.581
X4	Fertiliser	135	0.097	0.166	0	0.838
X5	Livestock	135	59.163	63.798	4.5298	280.18
Region 2 – The Americas						
Y	Output	105	11.698	26.443	0.484	127.350
X1	Land	105	58.536	102.170	1.343	431.400
X2	Machinery	105	330.540	1022.100	0.205	4800.000
X3	Labour	105	2.297	3.577	0.192	16.665
X4	Fertiliser	105	1.356	3.849	0.001	18.709
X5	Livestock	105	175.970	301.980	8.833	1253.500
Region 3 – Asia						
Y	Output	130	15.156	31.011	0.491	160.230
X1	Land	130	64.368	127.760	0.477	496.560
X2	Machinery	130	210.060	433.280	0.800	2142.200
X3	Labour	130	35.391	98.496	0.076	494.220
X4	Fertiliser	130	1.934	4.843	0	27.027
X5	Livestock	130	217.950	493.050	1.785	2389.800
Region 4 – Europe						
Y	Output	115	11.997	19.044	1.074	96.562
X1	Land	115	33.751	112.360	0.969	558.420
X2	Machinery	115	561.020	670.800	49.400	2780.000
X3	Labour	115	2.242	5.414	0.112	27.175
X4	Fertiliser	115	2.442	5.148	0.168	27.403
X5	Livestock	115	110.740	237.190	10.502	1208.200
All Countries						
Y	Output	485	10.077	22.772	0.484	160.230
X1	Land	485	46.353	100.170	0.477	558.420
X2	Machinery	485	265.590	648.490	0.085	4800.000
X3	Labour	485	11.930	52.949	0.076	494.220
X4	Fertiliser	485	1.418	4.061	0	27.403
X5	Livestock	485	139.240	320.470	1.785	2389.800

Table 3: Technical Efficiency and Technology Gap Ratio Estimates [a]

Region	Year			Mean	St.Dev	Minimum	Maximum
1	1986	TE	SFA-REG	0.499	0.253	0.200	0.971
			SFA-MF	0.365	0.176	0.155	0.868
			SFA-POOL	0.386	0.191	0.162	0.968
			DEA-REG	0.679	0.213	0.336	1.000
			DEA-MF	0.609	0.239	0.248	1.000
		TGR	SFA-TGR	0.772	0.193	0.355	1.000
			DEA-TGR	0.886	0.145	0.522	1.000
1	1987	TE	SFA-REG	0.500	0.252	0.202	0.971
			SFA-MF	0.360	0.172	0.158	0.848
			SFA-POOL	0.390	0.190	0.165	0.968
			DEA-REG	0.668	0.232	0.284	1.000
			DEA-MF	0.598	0.244	0.213	1.000
		TGR	SFA-TGR	0.765	0.201	0.299	0.992
			DEA-TGR	0.888	0.140	0.528	1.000
1	1988	TE	SFA-REG	0.502	0.252	0.204	0.971
			SFA-MF	0.363	0.170	0.149	0.813
			SFA-POOL	0.393	0.189	0.168	0.968
			DEA-REG	0.678	0.217	0.303	1.000
			DEA-MF	0.611	0.237	0.261	1.000
		TGR	SFA-TGR	0.769	0.204	0.340	1.000
			DEA-TGR	0.893	0.133	0.541	1.000
1	1989	TE	SFA-REG	0.503	0.251	0.205	0.971
			SFA-MF	0.365	0.173	0.144	0.814
			SFA-POOL	0.396	0.189	0.171	0.969
			DEA-REG	0.695	0.239	0.246	1.000
			DEA-MF	0.625	0.250	0.219	1.000
		TGR	SFA-TGR	0.772	0.216	0.332	1.000
			DEA-TGR	0.894	0.131	0.560	1.000
1	1990	TE	SFA-REG	0.504	0.251	0.207	0.971
			SFA-MF	0.366	0.175	0.147	0.820
			SFA-POOL	0.400	0.188	0.174	0.969
			DEA-REG	0.679	0.241	0.312	1.000
			DEA-MF	0.609	0.244	0.226	1.000
		TGR	SFA-TGR	0.773	0.219	0.262	1.000
			DEA-TGR	0.894	0.133	0.537	1.000
2	1986	TE	SFA-REG	0.751	0.170	0.504	0.970
			SFA-MF	0.559	0.164	0.336	0.970
			SFA-POOL	0.506	0.192	0.295	0.961
			DEA-REG	0.853	0.134	0.624	1.000
			DEA-MF	0.745	0.173	0.462	1.000
		TGR	SFA-TGR	0.754	0.166	0.423	1.000
			DEA-TGR	0.874	0.140	0.486	1.000
2	1987	TE	SFA-REG	0.767	0.160	0.532	0.972
			SFA-MF	0.566	0.159	0.341	0.934
			SFA-POOL	0.509	0.191	0.298	0.961
			DEA-REG	0.848	0.123	0.622	0.998
			DEA-MF	0.738	0.167	0.463	0.969
		TGR	SFA-TGR	0.746	0.164	0.404	0.974
			DEA-TGR	0.870	0.142	0.494	1.000

Table 3 continued

Region	Year			Mean	St.Dev	Minimum	Maximum
2	1988	TE	SFA-REG	0.782	0.151	0.559	0.974
			SFA-MF	0.573	0.158	0.346	0.911
			SFA-POOL	0.512	0.190	0.302	0.962
			DEA-REG	0.868	0.122	0.606	1.000
			DEA-MF	0.758	0.171	0.496	1.000
		TGR	SFA-TGR	0.739	0.162	0.394	0.977
			DEA-TGR	0.874	0.145	0.498	1.000
2	1989	TE	SFA-REG	0.797	0.142	0.586	0.976
			SFA-MF	0.586	0.156	0.342	0.903
			SFA-POOL	0.515	0.189	0.306	0.962
			DEA-REG	0.883	0.119	0.642	1.000
			DEA-MF	0.770	0.165	0.490	1.000
		TGR	SFA-TGR	0.743	0.168	0.353	0.991
			DEA-TGR	0.874	0.145	0.490	1.000
2	1990	TE	SFA-REG	0.810	0.133	0.611	0.978
			SFA-MF	0.598	0.153	0.371	0.913
			SFA-POOL	0.518	0.188	0.310	0.962
			DEA-REG	0.920	0.125	0.649	1.000
			DEA-MF	0.814	0.191	0.492	1.000
		TGR	SFA-TGR	0.745	0.168	0.397	1.000
			DEA-TGR	0.882	0.154	0.492	1.000
3	1986	TE	SFA-REG	0.706	0.194	0.348	0.981
			SFA-MF	0.537	0.208	0.224	0.944
			SFA-POOL	0.538	0.226	0.179	0.985
			DEA-REG	0.714	0.259	0.309	1.000
			DEA-MF	0.668	0.258	0.304	1.000
		TGR	SFA-TGR	0.758	0.183	0.391	0.995
			DEA-TGR	0.937	0.103	0.581	1.000
3	1987	TE	SFA-REG	0.707	0.198	0.351	0.981
			SFA-MF	0.526	0.209	0.226	0.949
			SFA-POOL	0.518	0.229	0.182	0.985
			DEA-REG	0.702	0.253	0.307	1.000
			DEA-MF	0.656	0.255	0.304	1.000
		TGR	SFA-TGR	0.745	0.193	0.377	1.000
			DEA-TGR	0.935	0.964	0.617	1.000
3	1988	TE	SFA-REG	0.709	0.197	0.353	0.981
			SFA-MF	0.531	0.206	0.230	0.931
			SFA-POOL	0.521	0.229	0.185	0.985
			DEA-REG	0.715	0.249	0.305	1.000
			DEA-MF	0.672	0.246	0.302	1.000
		TGR	SFA-TGR	0.752	0.195	0.376	0.986
			DEA-TGR	0.943	0.949	0.564	1.000
3	1989	TE	SFA-REG	0.710	0.196	0.356	0.982
			SFA-MF	0.534	0.209	0.233	0.904
			SFA-POOL	0.524	0.228	0.188	0.985
			DEA-REG	0.712	0.262	0.307	1.000
			DEA-MF	0.661	0.250	0.305	1.000
		TGR	SFA-TGR	0.753	0.199	0.375	1.000
			DEA-TGR	0.937	0.112	0.549	1.000

Table 3 continued

Region	Year			Mean	St.Dev	Minimum	Maximum
3	1990	TE	SFA-REG	0.712	0.195	0.359	0.982
			SFA-MF	0.534	0.208	0.236	0.894
			SFA-POOL	0.527	0.227	0.192	0.985
			DEA-REG	0.720	0.253	0.323	1.000
			DEA-MF	0.667	0.243	0.322	1.000
		TGR	SFA-TGR	0.752	0.203	0.363	1.000
			DEA-TGR	0.935	0.116	0.547	1.000
4	1986	TE	SFA-REG	0.756	0.154	0.451	0.980
			SFA-MF	0.537	0.190	0.240	0.959
			SFA-POOL	0.579	0.190	0.304	0.966
			DEA-REG	0.747	0.177	0.395	0.997
			DEA-MF	0.630	0.206	0.326	0.997
		TGR	SFA-TGR	0.716	0.210	0.297	1.000
			DEA-TGR	0.839	0.157	0.522	1.000
4	1987	TE	SFA-REG	0.760	0.152	0.459	0.981
			SFA-MF	0.537	0.190	0.239	0.956
			SFA-POOL	0.582	0.189	0.308	0.966
			DEA-REG	0.741	0.185	0.395	1.000
			DEA-MF	0.620	0.208	0.315	0.987
		TGR	SFA-TGR	0.712	0.211	0.295	0.993
			DEA-TGR	0.834	0.162	0.516	1.000
4	1988	TE	SFA-REG	0.764	0.150	0.468	0.981
			SFA-MF	0.537	0.191	0.241	0.928
			SFA-POOL	0.585	0.189	0.311	0.966
			DEA-REG	0.747	0.189	0.392	1.000
			DEA-MF	0.629	0.217	0.320	1.000
		TGR	SFA-TGR	0.706	0.210	0.289	0.987
			DEA-TGR	0.837	0.160	0.517	1.000
4	1989	TE	SFA-REG	0.769	0.147	0.476	0.982
			SFA-MF	0.543	0.196	0.243	0.953
			SFA-POOL	0.588	0.188	0.315	0.967
			DEA-REG	0.772	0.178	0.439	1.000
			DEA-MF	0.653	0.208	0.327	1.000
		TGR	SFA-TGR	0.710	0.214	0.291	0.995
			DEA-TGR	0.842	0.153	0.525	1.000
4	1990	TE	SFA-REG	0.773	0.145	0.484	0.982
			SFA-MF	0.558	0.202	0.260	0.977
			SFA-POOL	0.591	0.187	0.319	0.967
			DEA-REG	0.789	0.163	0.485	1.000
			DEA-MF	0.672	0.195	0.424	1.000
		TGR	SFA-TGR	0.726	0.223	0.300	1.000
			DEA-TGR	0.849	0.136	0.589	1.000
1	1986-1990	TE	SFA-REG	0.502	0.248	0.200	0.971
			SFA-MF	0.364	0.171	0.144	0.868
			SFA-POOL	0.393	0.187	0.162	0.969
			DEA-REG	0.680	0.225	0.246	1.000
			DEA-MF	0.610	0.239	0.213	1.000
		TGR	SFA-TGR	0.771	0.204	0.262	1.000
			DEA-TGR	0.891	0.134	0.522	1.000

Table 3 continued

Region	Year			Mean	St.Dev	Minimum	Maximum
2	1986-1990	TE	SFA-REG	0.781	0.150	0.504	0.978
			SFA-MF	0.577	0.156	0.336	0.970
			SFA-POOL	0.512	0.187	0.295	0.962
			DEA-REG	0.874	0.125	0.606	1.000
			DEA-MF	0.765	0.172	0.462	1.000
		TGR	SFA-TGR	0.745	0.162	0.353	1.000
			DEA-TGR	0.875	0.143	0.486	1.000
3	1986-1990	TE	SFA-REG	0.709	0.193	0.348	0.982
			SFA-MF	0.532	0.205	0.224	0.949
			SFA-POOL	0.526	0.224	0.179	0.985
			DEA-REG	0.713	0.251	0.305	1.000
			DEA-MF	0.665	0.247	0.302	1.000
		TGR	SFA-TGR	0.752	0.192	0.363	1.000
			DEA-TGR	0.937	0.103	0.547	1.000
4	1986-1990	TE	SFA-REG	0.764	0.147	0.451	0.982
			SFA-MF	0.543	0.190	0.239	0.977
			SFA-POOL	0.585	0.185	0.304	0.967
			DEA-REG	0.759	0.176	0.392	1.000
			DEA-MF	0.641	0.204	0.315	1.000
		TGR	SFA-TGR	0.714	0.210	0.289	1.000
			DEA-TGR	0.840	0.151	0.516	1.000
All	1986-1990	TE	SFA-REG	0.680	0.224	0.200	0.982
			SFA-MF	0.497	0.200	0.144	0.977
			SFA-POOL	0.500	0.209	0.162	0.985
			DEA-REG	0.750	0.216	0.246	1.000
			DEA-MF	0.666	0.227	0.213	1.000
		TGR	SFA-TGR	0.747	0.194	0.262	1.000
			DEA-TGR	0.888	0.137	0.486	1.000

[a] The DEA frontier was constructed using all observations (countries) in the data set. However, for purposes of comparison, the summary statistics reported in this table were calculated using only those observations (countries) that were available to calculate the SFA estimates.

Table 4: Technical Efficiency and Technology Gap Ratio Estimates for Selected Countries

Country	Year	TE					TGR	
		SFA-REG	SFA-MF	SFA-POOL	DEA-REG	DEA-MF	SFA-TGR	DEA-TGR
Argentina	1986	0.970	0.970	0.944	1	1	1	1
	1987	0.972	0.934	0.944	0.955	0.950	0.960	0.995
	1988	0.974	0.911	0.945	0.965	0.960	0.935	0.995
	1989	0.976	0.903	0.945	0.902	0.902	0.925	1
	1990	0.978	0.913	0.946	0.966	0.962	0.934	0.996
Australia	1986	0.949	0.944	0.761	0.933	0.919	0.995	0.985
	1987	0.949	0.949	0.763	0.941	0.909	1	0.966
	1988	0.950	0.931	0.765	0.969	0.930	0.981	0.960
	1989	0.950	0.902	0.767	0.904	0.874	0.949	0.967
	1990	0.950	0.872	0.769	0.931	0.913	0.918	0.981
Canada	1986	0.715	0.622	0.589	0.980	0.838	0.870	0.855
	1987	0.735	0.637	0.592	0.970	0.834	0.867	0.860
	1988	0.753	0.631	0.595	0.841	0.752	0.838	0.894
	1989	0.770	0.623	0.599	0.911	0.842	0.809	0.924
	1990	0.786	0.613	0.602	1	0.957	0.780	0.957
China	1986	0.937	0.895	0.607	0.309	0.304	0.955	0.984
	1987	0.937	0.886	0.610	0.307	0.304	0.946	0.990
	1988	0.938	0.886	0.614	0.305	0.302	0.945	0.990
	1989	0.938	0.886	0.617	0.307	0.305	0.945	0.993
	1990	0.939	0.890	0.620	0.323	0.322	0.948	0.997
India	1986	0.929	0.704	0.396	0.350	0.350	0.757	1
	1987	0.930	0.706	0.400	0.340	0.340	0.760	1
	1988	0.930	0.715	0.404	0.362	0.362	0.768	1
	1989	0.931	0.725	0.408	0.376	0.376	0.778	1
	1990	0.931	0.728	0.412	0.371	0.371	0.782	1
Indonesia	1986	0.587	0.558	0.781	0.935	0.791	0.952	0.846
	1987	0.589	0.569	0.783	0.823	0.755	0.966	0.917
	1988	0.591	0.583	0.785	0.763	0.734	0.986	0.962
	1989	0.594	0.592	0.787	0.713	0.697	0.997	0.978
	1990	0.596	0.596	0.789	0.685	0.676	1	0.987
Netherlands	1986	0.969	0.959	0.966	0.997	0.997	0.990	1
	1987	0.969	0.956	0.966	0.987	0.987	0.987	1
	1988	0.970	0.927	0.966	1	1	0.956	1
	1989	0.971	0.953	0.967	0.976	0.974	0.982	0.998
	1990	0.971	0.971	0.967	1	1	1	1
South Africa	1986	0.926	0.550	0.512	0.898	0.469	0.594	0.5222
	1987	0.926	0.560	0.516	0.978	0.516	0.605	0.528
	1988	0.926	0.551	0.519	0.955	0.517	0.595	0.541
	1989	0.927	0.556	0.523	1	0.560	0.600	0.560
	1990	0.927	0.560	0.526	0.984	0.528	0.604	0.537
UK	1986	0.963	0.573	0.590	0.745	0.737	0.595	0.989
	1987	0.964	0.579	0.593	0.747	0.738	0.601	0.988
	1988	0.965	0.579	0.597	0.732	0.724	0.601	0.989
	1989	0.965	0.568	0.600	0.741	0.732	0.589	0.988
	1990	0.966	0.566	0.603	0.761	0.749	0.586	0.984
USA	1986	0.957	0.807	0.961	0.999	0.968	0.843	0.969
	1987	0.961	0.815	0.961	0.978	0.959	0.848	0.981
	1988	0.964	0.826	0.962	0.906	0.890	0.857	0.982
	1989	0.967	0.847	0.962	0.961	0.960	0.876	0.999
	1990	0.969	0.854	0.962	1	1	0.881	1

Table 5: Estimates for Parameters of the Stochastic Frontier Models [a]

	Region 1 Frontier	Region 2 Frontier	Region 3 Frontier	Region 4 Frontier	Pooled Frontier	Meta-Frontier
β_0	-0.2 (1.4)	1.27 (0.93)	2.27 (0.79)	0.4 (2.3)	0.90 (0.29)	2.122
β_1	1.03 (0.54)	-0.76 (0.42)	0.08 (0.22)	1.9 (1.2)	0.06 (0.12)	-0.143
β_2	0.03 (0.24)	1.13 (0.36)	0.05 (0.20)	-1.15 (0.46)	0.123 (0.079)	0.010
β_3	0.15 (0.61)	0.44 (0.25)	0.96 (0.22)	-1.08 (0.61)	0.59 (0.11)	0.970
β_4	-0.04 (0.14)	0.15 (0.18)	0.50 (0.20)	0.32 (0.49)	0.167 (0.059)	0.338
β_5	-0.39 (0.48)	-0.72 (0.64)	-0.70 (0.36)	0.6 (1.0)	-0.15 (0.17)	-0.440
β_{11}	-0.36 (0.20)	-0.89 (0.22)	-0.140 (0.090)	0.43 (0.40)	-0.025 (0.044)	-0.066
β_{12}	0.075 (0.068)	0.267 (0.083)	0.038 (0.038)	0.10 (0.21)	0.070 (0.023)	0.036
β_{13}	0.06 (0.24)	0.16 (0.14)	-0.020 (0.054)	0.03 (0.27)	0.005 (0.035)	0.030
β_{14}	0.013 (0.028)	0.045 (0.047)	-0.054 (0.033)	-0.04 (0.12)	-0.018 (0.012)	-0.023
β_{15}	-0.08 (0.14)	0.65 (0.21)	0.052 (0.069)	-0.78 (0.17)	-0.077 (0.032)	0.026
β_{22}	0.083 (0.033)	0.05 (0.11)	-0.047 (0.037)	-0.22 (0.12)	0.033 (0.012)	0.048
β_{23}	0.22 (0.11)	-0.03 (0.11)	-0.202 (0.066)	-0.24 (0.11)	-0.041 (0.018)	-0.093
β_{24}	-0.014 (0.016)	-0.004 (0.053)	0.063 (0.039)	-0.021 (0.055)	-0.0035 (0.0080)	-0.010
β_{25}	-0.121 (0.069)	-0.45 (0.15)	0.136 (0.077)	0.59 (0.17)	-0.050 (0.029)	0.001
β_{33}	0.45 (0.39)	-0.18 (0.11)	0.233 (0.072)	-0.10 (0.20)	0.096 (0.044)	0.061
β_{34}	0.025 (0.051)	0.035 (0.048)	0.015 (0.034)	-0.051 (0.074)	0.015 (0.013)	0.025
β_{35}	-0.27 (0.27)	-0.148 (0.091)	-0.006 (0.074)	0.59 (0.13)	-0.074 (0.041)	-0.097
β_{44}	0.010 (0.016)	0.021 (0.026)	0.028 (0.018)	0.07 (0.19)	0.0249 (0.0056)	0.033
β_{45}	0.009 (0.027)	-0.035 (0.080)	-0.077 (0.061)	-0.03 (0.16)	0.006 (0.016)	-0.016
β_{55}	0.39 (0.25)	0.12 (0.28)	-0.05 (0.14)	-0.41 (0.30)	0.222 (0.081)	0.139
σ^2	0.86 (0.25)	0.078 (0.031)	0.229 (0.071)	0.115 (0.047)	0.77 (0.13)	
γ	0.9955 (0.0015)	0.9802 (0.0089)	0.9873 (0.0049)	0.9887 (0.0056)	0.99599 (0.00077)	
η	0.0049 (0.0064)	0.083 (0.012)	0.0075 (0.0088)	0.023 (0.012)	0.0103 (0.0022)	
Log-L	106.39	143.00	131.42	163.15	429.08	

[a] Standard errors are given in parentheses to two significant digits and the corresponding coefficient estimates are given to the same number of digits behind the decimal points as the standard errors.

Table 6: Peers

	Firm Identifier	Peer Count	Country	Year
Africa	4	23	Cameroon	1986
	6	3	Egypt	1986
	10	8	Cote Divore	1986
	19	67	Rwanda	1986
	33	2	Egypt	1987
	37	1	Cote Divore	1987
	46	9	Rwanda	1987
	64	1	Cote Divore	1988
	89	5	Ghana	1988
	100	12	Rwanda	1988
	102	14	Mali	1989
	112	26	Cameroon	1990
	114	13	Egypt	1990
	118	86	Cote Divore	1990
	132	22	Tunisia	1990
	133	32	Uganda	1990
The Americas	5	5	El Salvador	1986
	7	8	Haiti	1986
	12	38	Argentina	1986
	46	20	Dominican Republic	1988
	60	5	Paraguay	1988
	66	3	Cuba	1989
	67	22	Dominican Republic	1989
	70	2	Haiti	1989
	81	1	Paraguay	1989
	83	18	Uruguay	1989
	85	9	Canada	1990
	86	35	Costa Rica	1990
	87	1	Cuba	1990
	89	5	El Salvador	1990
	91	17	Haiti	1990
	92	2	Honduras	1990
	95	29	USA	1990
	97	5	Bolivia	1990
	99	22	Chile	1990
	101	29	Ecuador	1990
	102	2	Paraguay	1990
Asia	13	8	Laos	1986
	14	7	Malaysia	1986
	18	48	Philippines	1986
	25	1	New Zealand	1986
	26	5	PNG	1986
	36	2	Japan	1987
	44	10	Philippines	1987
	52	26	PNG	1987
	63	5	Cambodia	1988
	64	9	Korea Rep	1988
	66	25	Malaysia	1988
	77	37	New Zealand	1988
	78	14	PNG	1988
	88	1	Japan	1989
	90	6	Korea Rep	1989
	91	2	Laos	1989
	92	2	Malaysia	1989
	103	1	New Zealand	1989
	104	3	PNG	1989
	105	6	Bangladesh	1990
	113	58	Israel	1990
	114	3	Japan	1990
	118	1	Malaysia	1990
	119	6	Mongolia	1990
	122	16	Philippines	1990
	130	14	PNG	1990

continued next page

Table 6 continued.

	Firm Identifier	Peer Count	Country	Year
Europe	35	39	Italy	1987
	56	14	Hungary	1988
	59	36	Netherlands	1988
	71	17	Bel-Lux	1989
	78	4	Greece	1989
	79	11	Hungary	1989
	81	28	Italy	1989
	97	28	Denmark	1990
	102	31	Hungary	1990
	105	59	Netherlands	1990