

# Alternative Efficiency Measures for Multiple-Output Production

Carmen Fernández

*Department of Mathematics and Statistics, Lancaster University,  
Lancaster, LA1 4YF, U.K.*

Gary Koop

*Department of Economics, University of Leicester,  
Leicester, LE1 7RH, U.K.*

Mark Steel\*

*Department of Statistics, University of Warwick,  
Coventry, CV4 7AL, U.K.*

August 2003

This paper has two main purposes. Firstly, we develop various ways of defining efficiency in the case of multiple-output production. Our framework extends a previous model by allowing for nonseparability of inputs and outputs. We also specifically consider the case where some of the outputs are undesirable, such as pollutants. We investigate how these efficiency definitions relate to one another and to other approaches proposed in the literature. Secondly, we examine the behavior of these definitions in two examples of practically relevant size and complexity. One of these involves banking and the other agricultural data. Our findings are basically encouraging. For a given efficiency definition, efficiency rankings are found to be informative, despite the considerable uncertainty in the inference on efficiencies. It is, however, important for the researcher to select an efficiency concept appropriate to the particular issue under study, since different efficiency definitions can lead to quite different conclusions.

**Keywords:** banking, Bayesian inference, dairy farms, pollution, productivity, separability

JEL classification: C11, D24

---

\*Corresponding author: Mark Steel, Institute of Mathematics and Statistics, University of Kent at Canterbury, Canterbury, CT2 7NF, U.K.; e-mail: M.F.Steel@kent.ac.uk; fax: +44-1227-82 7932. As of October 1, 2003: Department of Statistics, University of Warwick, Coventry, CV4 7AL, U.K. fax: +44-24-7652 4532.

# 1 Introduction

The evaluation of the efficiency of individual firms is of fundamental importance for policymaking in many areas of economics. Stochastic frontier models have been one of the most popular tools for carrying out such efficiency analyses. Numerous applications in the fields of finance (*e.g.* Hunt-McCool, Koh and Francis, 1996), banking (*e.g.* Adams, Berger and Sickles, 1999, Fernández, Koop and Steel, 2000), agriculture (*e.g.* Kumbhakar, Ghosh and McGuckin, 1991), environmental economics (*e.g.* Reinhard, Lovell and Thijssen, 1999), public sector economics (*e.g.* Perelman and Pestieau, 1994) and development economics (*e.g.* Pitt and Lee, 1981) testify to the importance economists in diverse applied fields place on efficiency measurement. However, researchers must be cautious when using firm-specific efficiency measures to rank firms or make statements about whether a firm is more or less efficient than others. The necessity for caution arises for two reasons. Firstly, firm-specific efficiency is typically hard to estimate and associated measures of uncertainty (*e.g.* confidence intervals or Bayesian posterior standard deviations) can be quite large. Merely looking at point estimates can potentially be very misleading. Secondly, there is not one unique definition of efficiency and a firm which is ranked as being very efficient using one definition could potentially be ranked quite differently using another. The latter problem is exacerbated in the case of multiple-output production which forms the basis of the present paper.

These considerations motivate the focus of this paper. In particular, we develop and discuss several definitions of efficiency for the case of multiple-output production. We allow for non-separability of inputs and outputs by making the elasticity of transformation between outputs a parametric function of the inputs. We also treat the case where some of the outputs produced might be undesirable by-products of the production process, such as pollution. We shall use Bayesian methods for making inference about firm-specific efficiency using these definitions and compare the resulting efficiency rankings and conclusions in the context of two applications. We consider the empirically relevant case where the researcher only has data on inputs and outputs. That is, data on input or output prices, costs or profits are not available. Of course, if some or all of these were available, the problems involved in multiple-output production would be greatly simplified. For instance, if a cost function could be estimated, then outputs could be included as explanatory variables with cost being the dependent variable. However, this would not allow the development and analysis of a variety of output-oriented efficiency measures, each of which can shed a slightly different and potentially insightful light on the data.

Bayesian inference produces exact finite-sample posterior and predictive distributions and fully takes parameter uncertainty into account. It was found in previous work (see *e.g.* Koop, Osiewalski and Steel, 1994, 1997) to be an excellent tool for inference on efficiencies in stochastic frontier models, allowing *e.g.* for economic regularity conditions to be imposed in a very simple way. See Kim and Schmidt (2000) for some empirical comparisons of Bayesian and classical approaches to efficiency measurement.

The first of the empirical applications is on the banking data of Adams *et al.* (1999), also used in Fernández *et al.* (2000). The second is the environmental application of Reinhard *et al.* (1999) and Fernández, Koop and Steel (2002). Our findings indicate that, given an efficiency definition, efficiency rankings do have some practical relevance, despite the considerable uncertainty in the inference on efficiencies. On the other hand, it is critical that we focus on the appropriate efficiency concept for the particular purpose at hand, since different efficiency definitions can lead to quite different conclusions, both in terms of rankings and absolute values. The paper is organized as follows: In the second section we introduce our multiple-output production model. The third section discusses the issue of efficiency measurement in this model. The fourth section discusses some related approaches in the literature. Section 5 presents our two applications and the sixth concludes. Details on the prior adopted and the Markov chain Monte Carlo (MCMC) sampler used to conduct inference are given in the Appendix.

## 2 The Model

The best-practice technology for producing a vector of outputs,  $y$ , from a vector of inputs,  $x$ , can be described using a transformation function:

$$f(y, x) = 0.$$

In this paper we shall assume the transformation function has the form:

$$g(y, x) = h(x),$$

with a particular class of functions  $g(y, x)$ . The left and right hand sides of this equation are referred to as the aggregate output and production frontier, respectively. In the present paper, we adopt a modification of the setup of Fernández *et al.* (2000), hereafter FKS. As we shall see, our particular choice for the transformation function allows freeing up the separability assumption of our previous work in an attractive and intuitive manner.

We consider a set of  $NT$  observations corresponding to outputs of  $N$  different firms (or agents) where firm  $i$  is observed for  $T_i$  time periods. Note that this accommodates an unbalanced panel. For a balanced panel, we will have  $T_i = T$ ,  $i = 1, \dots, N$ . Generally, however,  $T$  will be defined as  $(T_1 + \dots + T_N)/N$ . The output of firm  $i$  ( $i = 1, \dots, N$ ) at time  $t$  ( $t = 1, \dots, T_i$ ) is a  $p$ -dimensional vector  $y_{(i,t)} = (y_{(i,t,1)}, \dots, y_{(i,t,p)})' \in \mathfrak{R}_+^p$ , and we assume that the aggregate output can be expressed as:

$$g(y_{(i,t)}, x_{(i,t)}) = \left( \sum_{j=1}^p \alpha_j^{q(x_{(i,t)})} y_{(i,t,j)}^{q(x_{(i,t)})} \right)^{1/q(x_{(i,t)})}, \quad (1)$$

where  $\alpha = (\alpha_1, \dots, \alpha_p)'$  with  $\alpha_j \in (0, 1)$  for all  $j = 1, \dots, p$  and  $\sum_{j=1}^p \alpha_j = 1$ . In addition, we assume that  $q(x_{(i,t)}) > 1$  is a function of the inputs  $x_{(i,t)}$  corresponding to firm  $i$  at time  $t$  given by:

$$q(x_{(i,t)}) = (1 + \psi_0) \prod_{l=1}^m (1 + \psi_l)^{x_{(i,t,l)}}, \quad (2)$$

where  $\psi = (\psi_0, \dots, \psi_m)'$  is a parameter vector in  $\mathfrak{R}_+^{m+1}$  and  $x_{(i,t,l)}$  is the  $l$ th input corresponding to observation  $(i, t)$ . To avoid cluttering the notation, we will not explicitly indicate the dependence of the output aggregator  $g(\cdot, \cdot)$  in (1) on  $\alpha$  and  $q(x_{(i,t)})$ . For fixed values of  $\alpha$  and  $q(x_{(i,t)})$ ,  $g(y_{(i,t)}, x_{(i,t)}) = \text{constant}$  defines a production equivalence surface, *i.e.* a  $(p - 1)$ -dimensional surface in  $\mathfrak{R}_+^p$  consisting of all output vectors  $y_{(i,t)}$  that are technologically equivalent. The aggregation function used in (1) has been used in Powell and Gruen (1968) and Kumbhakar (1987) and is sometimes called the ‘‘constant elasticity of transformation’’ aggregator, since it imposes the same elasticity of transformation between any two outputs. This elasticity of transformation is given by  $\{1 - q(x_{(i,t)})\}^{-1}$ , which explains the restriction  $q(x_{(i,t)}) > 1$  as this ensures a negative elasticity of transformation and also shows that this elasticity can be influenced by the input values. Thus, we have not imposed separability between inputs and outputs, and it is exactly through the most interesting property of the aggregator function that we allow for  $x_{(i,t)}$  to intervene. The interpretation of the parameter  $\alpha$  in (1) is to deal with scaling of the outputs, which is, therefore, quite separate from the role of  $q(x_{(i,t)})$ . Figure 1a illustrates the latter by displaying the range of production equivalence surfaces corresponding to various values of  $q(x_{(i,t)})$ . With given inputs in  $h(x_{(i,t)})$ , the production possibility set clearly increases with the value of  $q(x_{(i,t)})$ . In the special case of separability where  $q(x_{(i,t)}) = q$  and given  $q$ , the regularity conditions on  $h(x_{(i,t)})$  would imply a reduction of the production possibility set for smaller inputs  $x_{(i,t)}$  (monotonicity). If in our more general separable framework we would allow for  $q(x_{(i,t)})$  to be strictly decreasing in  $x_{(i,t)}$ , we could well face a situation where a reduction in input actually leads to a new produc-

tion possibility set which is not strictly contained in the old one<sup>1</sup>. In order to make sure that monotonicity in inputs holds for our nonseparable model, the specification in (2) imposes that  $q(x_{(i,t)})$  is a nondecreasing function of the inputs.

FKS use the transformation function in (1) with  $q(x_{(i,t)}) = q$ . The latter assumption corresponds to  $\psi_l = 0, l = 1, \dots, m$  and, thus, imposes the special case of separability in inputs and outputs of the production technology. As indicated by two Referees, this is a potentially restrictive assumption, so the present paper allows for nonseparability. We shall comment on the empirical support for this separability assumption in the context of our two applications in Section 5. In these applications, we shall use the levels of all inputs in (2), standardized to be in the interval  $[0, 1]$  by subtracting the maximum value in the sample. The latter helps in interpreting values for the coefficients  $\psi_1, \dots, \psi_m$  and makes a common prior (see (A.7) in the Appendix) a reasonable assumption.

Since the aggregate output  $g(y_{(i,t)}, x_{(i,t)})$  is a univariate quantity, it is natural to model it using a single-output stochastic frontier specification. To this end we define  $\delta_{(i,t)} = \log(g(y_{(i,t)}, x_{(i,t)}))$ , group these transformed outputs in an  $NT$ -dimensional vector

$$\delta = (\delta_{(1,1)}, \delta_{(1,2)}, \dots, \delta_{(1,T_1)}, \dots, \delta_{(N,T_N)})', \quad (3)$$

and model  $\delta$  as

$$\delta = V\beta - \gamma + \varepsilon. \quad (4)$$

In the latter equation,  $V = (v(x_{(1,1)}), \dots, v(x_{(N,T_N)}))'$  denotes an  $NT \times k$  matrix of exogenous regressors, where  $v(x_{(i,t)})$  is a  $k$ -dimensional function of the inputs  $x_{(i,t)}$ . The particular choice of  $v(\cdot)$  defines the specification of the production frontier: *e.g.* the vector  $v(x_{(i,t)})$  contains a 1 and all logged inputs for a Cobb-Douglas technology, whereas a translog frontier also involves squares and cross products of these logs. The corresponding vector of regression coefficients is denoted by  $\beta \in \mathcal{B} \subseteq \mathfrak{R}^k$ . Often, theoretical considerations will lead to regularity conditions on  $\beta$ , which will restrict the parameter space  $\mathcal{B}$  to a subset of  $\mathfrak{R}^k$ , still  $k$ -dimensional and possibly depending on  $x$ . For instance, we typically want to ensure that the marginal products of inputs are positive. Such conditions are easy to impose through the MCMC sampler described in the Appendix.

Technological inefficiency is captured by the fact that firms may lie below the frontier, thus leading to a vector of deviations between (the log of) actual and maximum possible aggregate

---

<sup>1</sup>The production equivalence surface would shift towards the origin close to the axes in Figure 1a, due to the reduction in  $h(x_{(i,t)})$ , but this could be more than offset by an outward shift of the surface for roughly equal values of both outputs, due to an increase in  $q(x_{(i,t)})$ .

output. This vector of deviations is labelled  $\gamma$ . It is usually reasonable to place some additional structure on these deviations. In particular, we set  $\gamma \equiv Dz \in \mathfrak{R}_+^{NT}$ , where  $D$  is a fixed  $NT \times M$  ( $M \leq NT$ ) matrix and  $z \in \mathcal{Z}$  with  $\mathcal{Z} = \{z = (z_1, \dots, z_M)' \in \mathfrak{R}^M : Dz \in \mathfrak{R}_+^{NT}\}$ . Through different choices of  $D$ , we can accommodate various amounts of structure on the vector  $\gamma$  of inefficiencies. For instance, taking  $D = I_{NT}$ , the  $NT$ -dimensional identity matrix, leads to an inefficiency term which is specific to each different firm and time period. For a balanced panel, choosing  $D = I_N \otimes \iota_T$ , where  $\iota_T$  is a  $T$ -dimensional vector of ones and  $\otimes$  denotes the Kronecker product, implies inefficiency terms which are specific to each firm, but constant over time (*i.e.* “individual effects”). In our empirical section we make the latter choice for  $D$ , but with the obvious extension to unbalanced panels for our second application. Fernández, Osiewalski and Steel (1997) provides a detailed description of other possible choices for  $D$ . In the next section, we will provide definitions of efficiency which will all be functions of this deviation  $\gamma$ . In previous work (see van den Broeck, Koop, Osiewalski and Steel, 1994), it was found that a reasonable choice for the distribution of  $z$  is a product of conditionally independent Exponentials. See the Appendix for more details.

The model in (4) also includes a two-sided error term  $\varepsilon$  which captures the fact that the frontier is not known exactly, but needs to be estimated from the data. We assume that  $\varepsilon$  follows an  $NT$ -dimensional Normal distribution with zero mean and covariance matrix equal to  $\sigma^2 I_{NT}$ . Hence, we obtain:

$$p(\delta|\beta, z, \sigma) = f_N^{NT}(\delta|V\beta - Dz, \sigma^2 I_{NT}), \quad (5)$$

where  $f_N^{NT}(\cdot|a, A)$  denotes the  $NT$ -variate Normal density function with mean  $a$  and covariance matrix  $A$ .

As discussed in FKS, the previous assumptions are not enough to specify the likelihood function. Intuitively, a single equation as in (5) is not sufficient to determine a probability density function for the  $p$ -dimensional vector  $y_{(i,t)}$  if  $p > 1$ . Hence, stochastics for the  $p - 1$  remaining dimensions must be specified by considering the distribution of the outputs within each of the production equivalence surfaces. Defining

$$\eta_{(i,t,j)} = \frac{\alpha_j^{q(x_{(i,t)})} y_{(i,t,j)}^{q(x_{(i,t)})}}{\sum_{l=1}^p \alpha_l^{q(x_{(i,t)})} y_{(i,t,l)}^{q(x_{(i,t)})}}, \quad j = 1, \dots, p, \quad \text{and} \quad \eta_{(i,t)} = (\eta_{(i,t,1)}, \dots, \eta_{(i,t,p)})', \quad (6)$$

we assume independent sampling across observational units (for  $i = 1, \dots, N$ ;  $t = 1, \dots, T_i$ ) from

$$p(\eta_{(i,t)}|s) = f_D^{p-1}(\eta_{(i,t)}|s), \quad (7)$$

where  $s = (s_1, \dots, s_p)' \in \mathfrak{R}_+^p$  and  $f_D^{p-1}(\eta|s)$  denotes the p.d.f. of a  $(p-1)$ -dimensional Dirichlet distribution with parameter  $s$  (see Poirier, 1995, page 132). From (6) we see that  $\eta_{(i,t)}$  can loosely be interpreted as a vector of output shares, with  $0 \leq \eta_{(i,t,j)} \leq 1$  and  $\sum_{j=1}^p \eta_{(i,t,j)} = 1$ . The Dirichlet distribution is a very flexible distribution commonly used to model shares.

The Appendix provides details regarding the prior distribution for all the parameters in the model. The posterior obtained from combining the prior with the likelihood function is not analytically tractable, but an MCMC algorithm can be developed to produce random draws from the posterior. These draws can then be used to obtain posterior properties for all of the efficiency measures. Details of the MCMC sampler are also briefly described in the Appendix.

### 3 Efficiency Measures

The previous section outlined an econometric model for multiple-output production where separability is not imposed and inefficiency is possible. If we focus on a single observation, and drop the  $(i, t)$  subscript from all variables for convenience, equation (4) can be written as:

$$g(y, x) = h(x)e^{-\gamma}e^\varepsilon, \quad (8)$$

where  $h(x)$  is the production frontier [*i.e.*  $\exp\{v(x)'\beta\}$ ]. Note that we do not explicitly indicate the dependence of  $g(\cdot, \cdot)$  on  $\alpha$  and  $q(x)$ , and of  $h(\cdot)$  on  $\beta$ . In order to derive explicit measures of efficiency, we relate the model in (8) to various efficiency concepts.

#### 3.1 Radial Output-Oriented Efficiency

The most common measure of efficiency is known as radial output-oriented technical efficiency,  $\tau_{RO}$ . The inverse of this (*i.e.*  $1/\tau_{RO}$ ) measures the amount by which all  $p$  outputs must be proportionally increased in order to get to the frontier. Clearly, for an inefficient firm we obtain  $\tau_{RO} < 1$ , whereas  $\tau_{RO} = 1$  corresponds to full efficiency.

The (stochastic) frontier faced by the firm with inputs  $x$  is given by  $h(x)e^\varepsilon$ . Hence, radial output-oriented efficiency is defined by the equation

$$g\left(\frac{y_1}{\tau_{RO}}, \dots, \frac{y_p}{\tau_{RO}}, x\right) = h(x)e^\varepsilon, \quad (9)$$

where  $0 < \tau_{RO} \leq 1$  and is equal to the output-distance function (see Shephard, 1970, or Färe and Primont, 1995) mentioned in some detail in Section 4. We can use (8) and (9) and the homogeneity of  $g(y, x)$  in (1) to obtain:

$$\tau_{RO} = e^{-\gamma}. \quad (10)$$

This is the definition of efficiency used in FKS. Note that the individual effects structure of  $\gamma$  will be inherited by  $\tau_{RO}$ , so that each firm will be assumed to have a specific radial output-oriented efficiency constant over time.

Radial output-oriented efficiency is illustrated in Figure 1b. This figure plots the production equivalence surface corresponding to the frontier (for a given level of inputs) for the two output case with  $q(x) = 1.5$ . In other words, for a given level of inputs, the possible output combinations that are achievable by a fully efficient firm are mapped out. If a firm using those inputs is producing at point A, its radial output-oriented efficiency is given by  $\tau_{RO} = |OA|/|OB|$ .

### 3.2 Individual Output-Oriented Efficiencies

In some cases, radial output-oriented efficiency, with its focus on a proportionate increase in all outputs, may not be the sensible efficiency measure. For instance, a bank may be interested in how much it can increase real estate loans, without affecting output of commercial loans. In this case, interest centers on the potential increase in one particular output and measures of individual output-oriented efficiencies, say,  $\tau_{jO}$  for  $j = 1, \dots, p$ , are called for. Formally,  $1/\tau_{jO}$  measures the amount by which output  $j$  must proportionally be increased in order to get to the frontier. Reasoning as in the previous subsection,  $\tau_{jO}$  is thus defined by:

$$g\left(y_1, \dots, \frac{y_j}{\tau_{jO}}, \dots, y_p, x\right) = h(x)e^\varepsilon. \quad (11)$$

Combining this equation with (8) and the form we use for  $g(\cdot, \cdot)$  (see equation 2.1), we obtain:

$$\tau_{jO} = e^{-\gamma} \left( \frac{\eta_j}{e^{-\gamma q(x)} \eta_j + 1 - e^{-\gamma q(x)}} \right)^{1/q(x)} = \tau_{RO} \left( \frac{\eta_j}{\tau_{RO}^{q(x)} \eta_j + 1 - \tau_{RO}^{q(x)}} \right)^{1/q(x)}, \quad (12)$$

where  $\eta_j$  is the “share” of output  $j$  described in (6), and where the second equality in (12) follows directly from the expression for the radial output-oriented efficiency in the previous subsection. It is clear from (12) that  $0 \leq \tau_{jO} \leq \tau_{RO}$ , with  $\tau_{jO} = \tau_{RO}$  only if  $\eta_j = 1$  (*i.e.* only the  $j^{th}$  output is produced) or if  $\tau_{RO} = 1$  (in which case the output vector  $y$  is on the frontier and both efficiency measures are equal to 1). We remind the reader that  $\eta_j$  is observation-specific [see (6)] and, thus, the ranking of firms in terms of  $\tau_{RO}$  is not necessarily preserved in terms of  $\tau_{jO}$ . In addition,  $\tau_{jO}$  can vary over time for the same firm (unlike  $\tau_{RO}$ ).

Individual output-oriented efficiency is illustrated in Figure 1b for each of the two outputs. In particular,  $\tau_{1O} = |O_1A|/|O_1B_1|$  and  $\tau_{2O} = |O_2A|/|O_2B_2|$ . Measures similar to these are used in Reinhard *et al.* (1999).

### 3.3 Efficiency Shares

The essential problem of efficiency analysis with multiple outputs can be seen in Figure 1b. There are many ways to measure the deviation of a firm (*e.g.* point A) from the production equivalence surface. The previous measures either used the distance along a radian connecting the origin to point A ( $\tau_{RO}$ ) or along a line parallel to an axis and passing through point A ( $\tau_{1O}$  and  $\tau_{2O}$ ). Here, we will consider an alternative concept, so-called “efficiency shares”, which involves mapping distances along the axes. An examination of Figure 1b indicates that  $\tau_{1S} = |OO_2|/|OC_1|$  and  $\tau_{2S} = |OO_1|/|OC_2|$  are intuitively plausible measures of efficiency.

The efficiency share for the  $j^{th}$  output,  $\tau_{jS}$ , is formally defined by the equation

$$g\left(0, \dots, \frac{y_j}{\tau_{jS}}, \dots, 0, x\right) = h(x)e^\varepsilon, \quad (13)$$

which has as solution

$$\tau_{jS} = \tau_{RO}\eta_j^{1/q(x)}, \quad (14)$$

where  $\eta_j$  is the share of output  $j$  as defined in (6). There are two ways in which we can interpret this efficiency measure. First, by equation (13),  $1/\tau_{jS}$  is the proportional increase of output  $j$  that can be achieved by being fully efficient while entirely giving up production of other outputs. In addition, for the output aggregator in (1) we note that  $g(0, \dots, 0, y_j/\eta_j^{1/q(x)}, 0, \dots, 0, x) = g(y_1, y_2, \dots, y_p, x)$ , so that  $(1 - \eta_j^{1/q(x)}) \geq 0$  is the relative reduction of production of output  $j$  due to the fact that other outputs are also being produced. Thus,  $\tau_{jS}$  in (14) represents the share of the (overall) efficiency  $\tau_{RO}$  due to the fact that several outputs are being jointly produced. In many cases, these efficiency shares will be an interesting measure of efficiency. From (12) and (14) we can immediately see that

$$\tau_{jS} \leq \tau_{jO} \leq \tau_{RO}, \quad (15)$$

where equalities hold if  $\eta_j = 1$  (graphically illustrated in Figure 1b by considering a point A on either of the axes). In contrast to  $\tau_{jO}$ , the efficiency share measure  $\tau_{jS}$  does not reduce to 1 for  $\tau_{RO} = 1$ .

### 3.4 Input-oriented Efficiency Measures

The previous discussion focussed on output-oriented measures. These all related to the question “by how much can output potentially be increased (using available inputs) if full efficiency is achieved?”. However, efficiency can also be measured using input-oriented measures which relate to the question “by how much can inputs potentially be decreased (holding output constant) if full efficiency is achieved?”.

Radial input-oriented efficiency,  $\tau_{RI}$ , is the most common input-oriented measure. It measures the proportionate decrease in all inputs  $x_1, \dots, x_m$  which is consistent with full efficiency. We can define efficient production as:

$$g(y_1, \dots, y_p, \tau_{RI}x_1, \dots, \tau_{RI}x_m) = h(\tau_{RI}x_1, \dots, \tau_{RI}x_m) e^\varepsilon.$$

Generally, this can be solved for a value of  $\tau_{RI}$ , but there is no simple closed-form solution available. However, in the special case of separability, we can derive from (8) that  $\tau_{RI}$  must satisfy

$$\ln[h(\tau_{RI}x_1, \dots, \tau_{RI}x_m)] - \ln[h(x_1, \dots, x_m)] + \gamma = 0. \quad (16)$$

Note that if constant returns to scale exists (*i.e.*  $h(\cdot)$  is homogeneous of degree 1), then  $\tau_{RI} = \exp(-\gamma) = \tau_{RO}$ . In other words, radial input- and output-oriented efficiencies are identical. However, if returns to scale are non-constant, these measures can be different (see, *e.g.*, Atkinson and Cornwell, 1994). In our first application, we assume a Cobb-Douglas production frontier,

$$h(x) = e^{\beta_0} x_1^{\beta_1} \dots x_m^{\beta_m}, \quad (17)$$

in which case

$$\tau_{RI} = e^{-\gamma/(\beta_1+\beta_2+\dots+\beta_m)} = \tau_{RO}^{1/(\beta_1+\dots+\beta_m)}. \quad (18)$$

Individual input-oriented measures, which address the question “how much can input  $l$  be decreased without sacrificing output”, can be derived in an analogous manner by defining efficient production as:

$$g(y_1, \dots, y_p, x_1, \dots, \tau_{lI}x_l, \dots, x_m) = h(x_1, \dots, \tau_{lI}x_l, \dots, x_m) e^\varepsilon.$$

Again, this only leads to a nice analytical solution in the separable case. If, in addition, we assume a Cobb-Douglas production frontier, we obtain:

$$\tau_{lI} = e^{-\gamma/\beta_l} = \tau_{RO}^{1/\beta_l}, \quad (19)$$

for  $l = 1, \dots, m$ .

Thus, with a Cobb-Douglas production frontier, the input-oriented measures are all powers of the standard radial output-oriented efficiency measure, and these powers do not vary across observations. Hence, rankings of firms in terms of their efficiency will be the same using  $\tau_{RO}$ ,  $\tau_{RI}$  or  $\tau_{lI}$  for  $l = 1, \dots, m$ . Accordingly, we do not investigate input-oriented efficiencies in our first application, as we find that separability is favored by the data.

In our second application we use a translog production frontier:

$$\ln[h(x)] = \beta_0 + \sum_{l=1}^m \beta_l \ln(x_l) + \sum_{j=1}^m \sum_{l \leq j} \beta_{lj} \ln(x_l) \ln(x_j). \quad (20)$$

Under separability we need to solve the equation in (16) and we obtain for radial input-oriented efficiency:

$$\ln(\tau_{RI}) = \frac{-b + \sqrt{b^2 - 4\gamma \sum_{j=1}^m \sum_{l \leq j} \beta_{lj}}}{2 \sum_{j=1}^m \sum_{l \leq j} \beta_{lj}}, \quad (21)$$

where  $b = \sum_{l=1}^m \beta_l + \sum_{j=1}^m \sum_{l \leq j} \beta_{lj} \ln(x_l x_j)$  and existence of the solution requires that  $b^2 > 4\gamma \sum_{j=1}^m \sum_{l \leq j} \beta_{lj}$ . Again assuming separability, we obtain for input  $l$ -oriented efficiency:

$$\ln(\tau_{II}) = \frac{-e_l + \sqrt{e_l^2 - 4\gamma\beta_{ll}}}{2\beta_{ll}}, \quad (22)$$

where  $e_l = \beta_l + \sum_{j \leq l} \beta_{jl} \ln(x_j) + \sum_{j \geq l} \beta_{lj} \ln(x_j)$  and we need that  $e_l^2 > 4\gamma\beta_{ll}$ .

A couple of things are worth noting about the input-oriented efficiency measures. Firstly,  $e_l$  is the elasticity of the frontier with respect to input  $l$ . To ensure that aggregate output is increasing in input, we impose (through our prior, see Appendix) the regularity conditions that  $e_l \geq 0$  for  $l = 1, \dots, m$  for every firm in every time period. This ensures that all the input-oriented efficiency measures are less than or equal to 1.0. Secondly, there are a few occasions where (21) and (22) do not yield real solutions (*i.e.*  $e_l^2 - 4\gamma\beta_{ll} < 0$  or  $b^2 - 4\gamma \sum_{j=1}^m \sum_{l \leq j} \beta_{lj} < 0$  for some  $i, t$ ). In the few cases where there is no value of  $\tau_{RI}$  which solves (16), we instead use for  $\tau_{RI}$  the value which makes it as close to zero as possible. It can be verified that this implies:  $\ln(\tau_{RI}) = -b/(2 \sum_{j=1}^m \sum_{l \leq j} \beta_{lj})$ . Reasoning in an analogous manner for  $\tau_{II}$  we obtain, for these cases,  $\ln(\tau_{II}) = -e_l/(2\beta_{ll})$ .

### 3.5 The Treatment of Good and Bad Outputs

In some applications, the jointness in production involves not only outputs that are desirable (so-called “good” outputs), but also some that are unavoidable by-products of the production process, such as pollution (which we will generally denote by “bad” outputs). In many empirical contexts, interest will be focused not only on how well firms do in producing good outputs, but also in how well they manage to avoid unnecessarily large production levels of the bad outputs. The question then becomes how to model such bad outputs. Earlier approaches were to include them as inputs (see *e.g.* Koop, 1998, and Reinhard *et al.*, 1999) or model them separately in a stochastic frontier context (Fernández *et al.*, 2002). The latter approach allows us to separately define two

types of efficiencies, “technical efficiency” related to the “goods” frontier and “environmental efficiency” corresponding to the “bads” frontier, which can both be estimated from the data. Here we are dealing with one single frontier and the issue of technical versus environmental efficiency involves defining different ways of looking at the same distance (between the frontier and the observed data). In this aspect, our present analysis is similar to that in Reinhard *et al.* (1999), where both efficiency measures are clearly related.

We now have, say,  $p_1 < p$  good outputs and  $p - p_1$  bad outputs, and reorder the output vector to have the good outputs as the first elements. In defining production equivalence surfaces, we need to take into account that a larger production of good outputs will also imply a larger production of bad outputs, so we can not directly use the output aggregator in (1). We can, however, still use the latter aggregator if we transform bads ( $y_j$ ) to negative powers ( $y_j^{-r}$ ). Thus, our output aggregator now becomes

$$g(y, x) = \left( \sum_{j=1}^{p_1} \alpha_j^{q(x)} y_j^{q(x)} + \sum_{j=p_1+1}^p \alpha_j^{q(x)} y_j^{-rq(x)} \right)^{1/q(x)}, \quad (23)$$

where the value of  $r > 0$  still remains to be chosen. Figure 2 plots the production equivalence surface defined through fixing  $g(y, x)$  in a case where we have one good output and one bad output. The production equivalence surfaces are plotted for  $q(x) = 1.1$  and various values of  $r$  (namely  $r = 0.5$ ,  $r = 1$  and  $r = 3$ ). Note from (23) and Figure 2 that it is not possible to produce goods without bads; this is in line with the idea that bads are unavoidable by-products of the goods production process and is implied by the definition of null-jointness in *e.g.* Färe, Grosskopf, Noh and Weber (2002). This new output aggregator inspires the following definitions.

Technical efficiency,  $\tau_T$ , is defined through

$$g\left(\frac{y_1}{\tau_T}, \dots, \frac{y_{p_1}}{\tau_T}, y_{p_1+1}, \dots, y_p, x\right) = h(x)e^\varepsilon, \quad (24)$$

which leads to

$$\tau_T = e^{-\gamma} \left( \frac{\eta_T}{e^{-\gamma q(x)} \eta_T + 1 - e^{-\gamma q(x)}} \right)^{1/q(x)}, \quad (25)$$

after defining

$$\eta_T = \frac{\sum_{j=1}^{p_1} \alpha_j^{q(x)} y_j^{q(x)}}{\sum_{j=1}^{p_1} \alpha_j^{q(x)} y_j^{q(x)} + \sum_{j=p_1+1}^p \alpha_j^{q(x)} y_j^{-rq(x)}}. \quad (26)$$

Note the similarity of (25) to the individual output-oriented efficiency in (12), where we have now used the share of good outputs in the aggregator function,  $\eta_T$ . Clearly,  $\tau_T \leq \exp(-\gamma)$

with equality only if  $\eta_T = 1$  (infinite bads) or if  $\gamma = 0$ . The case of infinite bads<sup>2</sup> can be interpreted as the situation where we focus entirely on producing goods and do not attempt to reduce bad outputs, so that the technical efficiency as measured by the goods only is equal to the radial output-oriented efficiency. In other cases, we sacrifice some technical efficiency as a consequence of attempts to reduce pollution, due to the trade-off implicit in our aggregator. If we focus on the case  $r = 3$  in Figure 2, a production level at point  $A$  will imply a technical efficiency equal to  $|O_2A|/|O_2B_2|$ .

Similarly, we can define environmental efficiency,  $\tau_E$ , through

$$g(y_1, \dots, y_{p_1}, \tau_E y_{p_1+1}, \dots, \tau_E y_p, x) = h(x)e^\varepsilon, \quad (27)$$

which, with  $g(y, x)$  as in (23), leads to

$$\tau_E = e^{-\gamma/r} \left( \frac{\eta_E}{e^{-\gamma q(x)} \eta_E + 1 - e^{-\gamma q(x)}} \right)^{\frac{1}{r q(x)}}, \quad (28)$$

where we have defined  $\eta_E = 1 - \eta_T$  as the relative contribution of the bads in the aggregator. In this case,  $\tau_E \leq \exp(-\gamma/r)$  with equality only if  $\eta_E = 1$  (no goods<sup>3</sup>) or if  $\gamma = 0$ . A graphical illustration of environmental efficiency is given in Figure 2, by realizing that an observation at  $A$  with  $r = 3$  induces  $\tau_E = |O_1B_1|/|O_1A|$ .

In the context of data envelopment analysis (DEA), Tyteca (1996) discusses a measure of environmental performance that is in the same spirit as the inverse of our  $\tau_E$ . Again in a DEA framework, Färe, Grosskopf and Zaim (1999) consider an environmental performance index, which is the ratio of a quantity index based on a measure akin to  $\tau_T$  and an index derived from the measure in Tyteca (1996) mentioned above.

## 4 Discussion of Related Work

There is an extensive literature which considers efficiency definitions closely related to those presented here (see, *e.g.*, Färe and Primont, 1995). Although the terminology and notation is different from ours, the basic ideas are quite similar. A key concept in this literature is the output distance function,  $D_0(x, y)$ , defined as:

---

<sup>2</sup>As pointed out by a Referee, the aggregator in (23) does formally allow for unbounded quantities of bad output with finite inputs, which is clearly not a literal reflection of reality. However, for empirically relevant combinations of good and bads, we feel the properties of our aggregator function are quite reasonable, and we would not want to use the model for predictions far outside the range of the data.

<sup>3</sup>As a Referee remarks, in the hypothetical case of no goods production, our aggregator function in (23) would imply a residual quantity of bads. Again, we consider our model to be reasonable in the empirically relevant range.

$$D_0(x, y) = \inf \left\{ \theta : \left( \frac{y}{\theta} \right) \in P(x) \right\},$$

where  $\theta$  is a scalar and  $P(x)$  is an output correspondence. That is,  $P(x)$  defines the set of outputs which can be produced by  $x$ . It can be seen that our equation (4), with the measurement error removed, defines an output distance function and  $\theta$  is equivalent to our radial output-oriented efficiency measure,  $\tau_{RO}$ , in (9). An input distance function can be defined analogously and related to our radial input-oriented efficiency.

Most of the work using distance functions has been implemented using non-econometric methods (*e.g.* Malmquist productivity indices are calculated using DEA techniques by, among many others, Färe, Grosskopf and Kirkley, 2000). Researchers who adopt econometric approaches include Atkinson, Cornwell and Honerkamp (2003) who use a single-equation Generalized Method of Moments methodology and Coelli and Perelman (2002) who use a single-equation corrected OLS methodology. However, for likelihood-based inference, such as our Bayesian analysis, we do need more than such a univariate distance function as we are required to specify a distribution in  $p$  dimensions.

The concept of the distance function has been extended in, *e.g.*, Chambers, Chung and Färe (1996) to the directional distance function. The output-oriented directional distance function is defined as:

$$\vec{D}_0(x, y, d) = \sup \{ \lambda : y + \lambda d \in P(x) \},$$

where  $d$  is a  $p \times 1$  direction vector and  $\lambda$  is a scalar. The relationship between the output distance function and output-oriented directional distance function is obtained by setting  $d = y$  and noting:

$$D_0(x, y) = \left\{ 1 + \vec{D}_0(x, y, y) \right\}^{-1}.$$

The directional distance function can be used to measure inefficiency in terms of distances from an observation to the frontier in any direction. Thus, the close relationship between  $\tau_{jO}$  and  $\vec{D}_0(x, y, d_j)$  with  $d_j = (0, \dots, 0, y_j, 0, \dots, 0)'$  is clear. Our efficiency shares also represent another direction in which distance to the frontier can be measured. Directional distance functions have typically been estimated using DEA techniques (see, *e.g.*, Ball, Färe, Grosskopf and Nehring, 2001) and input-oriented directional distance functions can be defined which are closely related to our individual input-oriented efficiency measures. In the context of goods and bad outputs, Färe *et al.* (2002) impose null-jointness and weak disposability of (good and bad) outputs and base efficiency measures on directional distance functions.

## 5 Applications

In this section, we investigate the properties of the various efficiency measures in the context of two substantive applications which have received attention in the applied literature. It is worth noting that in both applications  $q$  tends not to be far from one for most firms. This may help the reader is assessing some of the properties of the expressions in Section 3. Reported results on the banking data are based on a Monte Carlo run of 25,000 retained draws after discarding the first 5000, whereas results for the agricultural data correspond to a run of 60,000 and a burn-in of 10,000 draws.

### 5.1 Efficiency of the US Banking Industry

Berger and Humphrey (1997) provide an extensive survey of the literature on the efficiency of financial institutions. We use a data set which has been used by several others to investigate this important issue (see Berger, 1993, Adams *et al.*, 1999, and FKS). This banking data set contains observations on  $N = 798$  limited branching banks in the United States for  $T = 10$  years (1980-89). The theory of bank behavior suggests that it is reasonable to treat loans as an output in a production process where deposits and traditional factors (*e.g.* labor, capital) are inputs. Accordingly, the data set contains  $p = 3$  outputs (real estate loans; commercial and industrial loans; installment loans) and  $m = 5$  inputs (average number of employees; physical capital; purchased funds; demand deposits; retail, time and savings deposits). A Cobb-Douglas form is assumed for the production frontier. We assume that  $\gamma$  has an individual effects structure and, hence,  $D = I_N \otimes \iota_T$  (see the discussion following (4)).

FKS carries out a Bayesian analysis of this data set. The model used in FKS contains explanatory variables in the efficiency distribution, and imposes separability. With these exceptions, the present model (described in Section 2 and the Appendix) is identical to that in FKS and results are quite similar. We focus mainly on the differences with FKS and on efficiency measurement. Relative to FKS, we have extended our model to allow for nonseparability through (2) and, thus, it is worthwhile to present results for  $\psi$ .<sup>4</sup> Table 1 presents the medians of the posterior distributions as point estimates and 95% Bayesian credible intervals<sup>5</sup> for the  $m + 1$  elements of  $\psi$ . Since  $\psi_1, \dots, \psi_5$  have most of the posterior mass quite close to zero, there is no immediate evidence of a departure from separability of the transformation function (corresponding to  $\psi_l = 0$

---

<sup>4</sup>To aid in interpretation, note that the explanatory variables used for  $q$  in (2) have been standardized to take values in  $[0, 1]$ .

<sup>5</sup>Defined as ranging from the 2.5th percentile to the 97.5th percentile of the posterior distribution.

for  $l = 1, \dots, 5$ ). This is in line with the Bayes factor which mildly favors the separable model over the nonseparable one: the Bayes factor is estimated to be 40.4 in favor of the nonseparable model. This is based on the marginal likelihood values for each model, which are estimated using the modified harmonic mean estimator  $p_4$  of Newton and Raftery (1994). Thus, for this data set, there is no real evidence of nonseparability of the transformation function. However, we shall present the results of the more general nonseparable model in the sequel (which are quite close to those with the separable model). In addition, Table 1 tells us that the posterior for  $\psi_0$  is located very near to zero, indicating values of  $q(x) \approx 1$ . Thus, in this application, we are near the limiting case of a linear production equivalence surface (see Figure 1a for  $q(x) = 1$ ).

		95% Credible Interval	
	Median	Lower	Upper
$\psi_0$	$4.0 \times 10^{-5}$	$1.0 \times 10^{-7}$	$1.9 \times 10^{-4}$
$\psi_1$	$2.9 \times 10^{-3}$	$9.6 \times 10^{-5}$	$1.5 \times 10^{-2}$
$\psi_2$	$3.8 \times 10^{-3}$	$1.4 \times 10^{-4}$	$2.0 \times 10^{-2}$
$\psi_3$	$7.3 \times 10^{-3}$	$3.0 \times 10^{-4}$	$3.8 \times 10^{-2}$
$\psi_4$	$8.0 \times 10^{-5}$	$5.9 \times 10^{-6}$	$4.2 \times 10^{-4}$
$\psi_5$	$1.5 \times 10^{-4}$	$1.1 \times 10^{-5}$	$3.4 \times 10^{-4}$

We next investigate whether the use of various definitions of efficiency allows us to uncover potentially interesting aspects of the data that would remain hidden if we focussed only on, say,  $\tau_{RO}$ . We begin with a simple analysis based only on point estimates of efficiency. Note that strong assumptions are necessary to derive classical confidence intervals for firm-specific efficiencies (see Horrace and Schmidt, 1996) and common computer packages (*e.g.* LIMDEP) do not produce confidence intervals. It is perhaps for these reasons, that applied work often merely presents point estimates of efficiencies (*e.g.*, among many others, Barla and Perelman, 1989). We use posterior means as point estimates.

Given the individual effects structure of  $\gamma$ ,  $\tau_{RO}$  will also have an individual effects structure. However,  $\tau_{jO}$  and  $\tau_{jS}$  will not ( $j = 1, \dots, p$ ). Accordingly, for each of our  $2p + 1 = 7$  output-oriented efficiency definitions, we have a  $TN$ -vector which contains the posterior means of efficiency for each bank in each time period. Table 2 presents the correlation matrix between these seven vectors. This simple measure of linear interdependence is computed as the correlation coefficient between pairs of observations, where pairs share the same  $(i, t)$  index. Table 3 presents the Spearman rank correlation matrix computed along the same lines.

Consider first radial output-oriented efficiency. All of the other efficiency measures are positively correlated with this, as could be expected from the expressions in (12) and (14), but some

$\tau_{RO}$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{3O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{3S}$
1.00						
0.89	1.00					
0.71	0.52	1.00				
0.87	0.72	0.59	1.00			
0.61	0.86	0.15	0.34	1.00		
0.26	-0.02	0.75	0.04	-0.22	1.00	
0.66	0.44	0.36	0.90	0.03	-0.10	1.00

$\tau_{RO}$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{3O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{3S}$
1.00						
0.88	1.00					
0.71	0.47	1.00				
0.88	0.67	0.54	1.00			
0.59	0.88	0.17	0.35	1.00		
0.25	0.00	0.81	0.09	-0.18	1.00	
0.65	0.41	0.33	0.92	0.10	-0.04	1.00

of these correlations are quite low (*e.g.* the correlation with  $\tau_{2S}$ ). This provides preliminary evidence that alternative efficiency measures can provide a different perspective on the same data set. Between the efficiency shares we find correlations near zero. This is not so surprising since  $\tau_{jS} = \eta_j^{1/q(x)} \tau_{RO}$  [see (14), where  $q(x)$  takes values close to 1] and the correlation between  $\eta_i$  and  $\eta_j$  for  $i \neq j$  will be negative<sup>6</sup>. This implies a negative contribution to the correlation between  $\tau_{jS}$  and  $\tau_{iS}$ , which counteracts or even dominates the positive contribution through the common factor  $\tau_{RO}$ . Further, correlations between  $\tau_{iO}$  and  $\tau_{jS}$  tend to be quite high for  $i = j$  but considerably lower for  $i \neq j$ .

As discussed above, the input-oriented efficiency measures for the Cobb-Douglas production frontier are always the same powers of radial output-oriented efficiency in the separable case. Given that the data favor the separable model, and the results presented here are very close to those with that model, these powers will be quite an accurate approximation of the input-oriented measures and, hence, the latter are not presented here.

Note that results using the Spearman rank and simple correlations are quite similar to one another. We found this similarity in all of our results and, accordingly, do not discuss the Spearman rank correlation in our subsequent empirical work.

---

<sup>6</sup>This is not counterintuitive, since the  $\eta_j, j = 1, \dots, p$ , add up to unity.

In order to further investigate the relationships between the different efficiency measures, we created scatterplots from the point estimates of efficiency used to calculate the correlations in Table 2. For the sake of brevity, we present only selected results. Figures 3, 4 and 5 plot the point estimates of  $\tau_{RO}$  against  $\tau_{jS}$  for  $j = 1, 2, 3$ , respectively. As indicated in Subsection 3.3, we always have  $\tau_{RO} \geq \tau_{jS}$  and, hence, the points in the figures all lie below the 45-degree line. It is clear that there are many observations which lead to substantially different efficiency estimates using the different efficiency measures. The main conclusion we can draw from these figures is that indeed  $\tau_{2S}$  behaves quite differently from the usual radial output-oriented efficiency. In particular, the firms that are closest to the frontier (with  $\tau_{RO} > 0.8$ ) seem to produce relatively less of the second output (commercial and industrial loans) and shift mostly to the third output (installment loans). This conclusion was borne out by looking at scatterplots of posterior estimates of  $\tau_{RO}$  versus  $\eta$ . Thus, if we are particularly interested in how well firms do in terms of the second output (with respect to what they could produce if they gave up production of the other outputs), the conclusions in terms of efficiency levels and ranking of firms will inevitably be very different from those based on  $\tau_{RO}$ .

If we conduct a similar comparison of radial and individual output-oriented efficiencies, we find roughly similar results, although the differences are not that striking.

What should we make of our previous results? Clearly, we cannot say “It does not matter which efficiency definition you choose, all will give essentially the same results”. However, in a practical setting, the choice between different efficiency measures is not likely to be completely arbitrary. Researchers who are interested in departing from the standard radial-oriented output measures, will typically have a reason to focus on a particular output (*e.g.* to measure the shortfall of real estate loans from the maximum possible). Given the rather different interpretation of individual output-oriented efficiency and efficiency shares, it is not surprising that they would give different results for some observations. Thus, what this really establishes is that the researcher must make a choice between the various measures of output-oriented efficiency. In addition, the variety of efficiency measures we present allows us to extract and highlight more information from a given data set.

The numbers in the previous tables were obtained by first calculating the posterior mean for each firm and time period and then calculating correlations. This, of course, ignores the estimation uncertainty inherent in any efficiency analysis. A simple way to shed light on this is to treat the correlations between efficiency measures as random variables and plot their posteriors. This can be done by calculating the correlation of interest for each MCMC draw and then plotting

its posterior density function. This is done in Figures 6-8 for selected efficiency measures. For the sake of brevity, we do not plot the posteriors for all  $2p^2 + p$  correlations. Rather Figures 6 and 7 plot the posteriors for the correlations between  $\tau_{RO}$  and each of  $\tau_{jO}$  and  $\tau_{jS}$  for  $j = 1, \dots, p$ . Figure 8 plots the posterior of the correlation between the individual output efficiency and efficiency share measures for the same output (*i.e.* the correlation between  $\tau_{jO}$  and  $\tau_{jS}$ ). These figures do nothing to undermine our previous findings. That is, the correlations between the efficiency measures are fairly precisely estimated. Thus, our previous results based on point estimates point in the same direction as those which properly address the issue of estimation uncertainty.

Another way to consider the differences in efficiency measures is to monitor whether firms can make large jumps in the efficiency ranking if we use a different measure. In order to do this, we take the estimates of  $\tau_{RO}$  for each firm, rank them, and select the firms ranked at the first, second and third quartiles. We then calculate the posterior probability that the relative ranking of each pair of firms is maintained, using each of the efficiency measures.

Table 4 contains the results. If we begin by examining the column labelled  $\tau_{RO}$ , we can see that rankings based on point estimates do give a crude but reliable picture of firm-specific banking efficiency. That is, the firm ranked as being at the 3rd quartile is almost always more efficient than the median and 1st quartile firms. This finding indicates that if a researcher decides on one particular efficiency definition (*e.g.*  $\tau_{RO}$ ), then the large spread of efficiencies (evidenced in Figures 3-5) does not prevent at least an approximate classification of observed firms in terms of their mean efficiency levels.

Firm A/Firm B	$\tau_{RO}$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{3O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{3S}$
3rd Quart./Median	0.98	0.94	0.25	1.00	0.68	0.00	1.00
3rd Quart./1st Quart.	1.00	1.00	1.00	1.00	0.29	0.04	1.00
Median/1st Quart.	1.00	0.90	1.00	1.00	0.18	1.00	1.00

Three of the columns of Table 4 exhibit a similar pattern. In particular, using  $\tau_{1O}$ ,  $\tau_{3O}$  or  $\tau_{3S}$  as efficiency measures would lead to a similar differentiation between these three firms as the one based on  $\tau_{RO}$ . However, the other efficiency definitions indicate somewhat different rankings of the three firms. Tables 5 and 6 present posterior means and standard deviations, respectively, for the efficiencies of the three firms using all of the definitions. We know that the absolute magnitudes of the various efficiencies are subject to the inequalities in (15). An examination of these tables immediately puts the results in Table 4 in context. Posterior standard deviations of, say,  $\tau_{RO}$  are small enough to lend some practical relevance to the ranking of the three firms.

However, the three firms selected as being at the first, second and third quartiles using  $\tau_{RO}$  are ranked quite differently using some of the other efficiency definitions. In particular, the firm ranked as being quite efficient by  $\tau_{RO}$  (*i.e.* the third quartile), is quite inefficient according to  $\tau_{2S}$  and is below the median according to  $\tau_{2O}$ . In addition, all three firms have roughly equal values for  $\tau_{1S}$ . Thus, the results of Tables 5 and 6 highlight the need for a careful choice of the efficiency measure used, depending on the question one wants to address.

Ranking acc. to $\tau_{RO}$	$\tau_{RO}$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{3O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{3S}$
3rd Quartile	0.73	0.35	0.44	0.57	0.15	0.21	0.37
Median	0.63	0.28	0.47	0.31	0.14	0.33	0.16
1st Quartile	0.51	0.24	0.32	0.19	0.15	0.24	0.12

Ranking acc. to $\tau_{RO}$	$\tau_{RO}$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{3O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{3S}$
3rd Quartile	0.04	0.04	0.04	0.04	0.01	0.01	0.02
Median	0.04	0.03	0.04	0.03	0.01	0.02	0.01
1st Quartile	0.02	0.02	0.02	0.01	0.01	0.01	0.01

Incidentally, the posterior mean value for  $\tau_{RO}$  corresponding to the median firm accords well with the estimated average relative efficiencies computed for these data on the basis of non-parametric methods in Adams *et al.* (1999).

Researchers may worry about presenting firm-specific efficiency estimates or rankings on the grounds that posterior standard deviations might be large and, hence, point estimates might be misleading. Alternatively, they may worry about the choice of efficiency definition. Our findings above indicate that it is the latter issue that is the more serious one. Posterior standard deviations, at least for the present application, are not so large as to preclude at least a rough categorization of firms in terms of their efficiency. However, alternative efficiency definitions can give very different results. Thus, we should carefully choose the efficiency measure that is the most appropriate one to answer the question of relevance to us in a particular empirical situation.

## **5.2 An Application to Dutch Dairy Farms**

Stochastic frontier methods have been extensively used in the field of agricultural economics. As an example in this area, we use a data set compiled by the Agricultural Economics Research Institute in the Hague from highly specialized dairy farms that were in the Dutch Farm Accountancy Data Network, a stratified random sample. Reinhard *et al.* (1999) and Reinhard (1999)

describe the data in detail. The panel is unbalanced and we have 1,545 observations on  $N = 613$  dairy farms in the Netherlands for some or all of 1991-94. The dairy farms produce three outputs ( $p = 3$ ), one of which is bad ( $p_1 = 2$ ) using three inputs ( $m = 3$ ):

- Good Outputs: Milk (millions of kg) and Non-milk (millions of 1991 Guilders).
- Bad Output: Nitrogen surplus (thousands of kg).
- Inputs: Family labor (thousands of hours), Capital (millions of 1991 Guilders) and Variable input (thousands of 1991 Guilders).

Variable input includes *inter alia* hired labor, concentrates, roughage and fertilizer. Non-milk output contains meat, livestock and roughage sold. The definition of capital includes land, buildings, equipment and livestock.

Fernández *et al.* (2002) analyze the same data, using two separate Cobb-Douglas frontiers (one for the goods and one for the bad). Here, we assume a translog form for the single production frontier, as in Reinhard *et al.* (1999). Hence, using the notation following equation (4), and denoting the  $l$ th element of  $x_{(i,t)}$  by  $x_{(i,t,l)}$ , we have  $k = 10$  and:

$$v(x_{(i,t)}) = (1, \ln(x_{(i,t,1)}), \dots, \ln(x_{(i,t,3)}), \ln^2(x_{(i,t,1)}), \ln(x_{(i,t,1)}) \ln(x_{(i,t,2)}), \dots, \ln(x_{(i,t,2)}) \ln(x_{(i,t,3)}), \ln^2(x_{(i,t,3)}))'$$

We experimented with different values for  $r$  in the aggregator function (23) and found the results relatively insensitive to the actual choice of  $r$  in the range 0.5 to 3. We shall present our results for  $r = 3$ . Again, we assume an individual effects structure for  $\gamma$ .

As before, we focus mainly on the efficiencies. As in the case of the banking data, the data favor separability in the production frontier. Marginal likelihoods computed using the estimator  $p_4$  of Newton and Raftery (1994) lead to a Bayes factor of 63.2 in favor of the separable model. Thus, the evidence that the elasticity of transformation does not depend on the inputs used is even stronger for the farm data than for the banking data. Separability also greatly facilitates the calculation of input-oriented efficiencies (see Subsection 3.4), which we want to examine in this translog application. Therefore, we now only present results for the separable model.

Table 7 presents the correlations between the measures of efficiency introduced in Section 3, including the input-oriented measures and the technical and environmental efficiencies of Subsection 3.5. Results using Spearman rank correlations are very similar.

All input-oriented measures are highly correlated among themselves, and also display a high positive correlation with the output-oriented efficiency measures in the output 1 (milk) direction.

$\exp(-\gamma)$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{RI}$	$\tau_{1I}$	$\tau_{2I}$	$\tau_{3I}$	$\tau_T$	$\tau_E$
1.00										
0.98	1.00									
0.56	0.44	1.00								
0.91	0.97	0.23	1.00							
0.05	-0.11	0.72	-0.34	1.00						
0.98	0.97	0.52	0.91	0.04	1.00					
0.76	0.75	0.65	0.70	-0.01	0.70	1.00				
0.88	0.87	0.53	0.82	-0.02	0.85	0.81	1.00			
0.88	0.85	0.66	0.76	0.10	0.83	0.87	0.90	1.00		
1.00	0.98	0.56	0.92	0.06	0.98	0.75	0.87	0.87	1.00	
-0.46	-0.50	-0.18	-0.50	-0.04	-0.61	-0.18	-0.39	-0.32	-0.51	1.00

For output 2-oriented measures this is not the case (especially for efficiency shares). Clearly, efficiency measurement corresponding to output 2 (non-milk) is not very similar to the other measures. Technical efficiency [as defined in (25)] is highly positively correlated with all other measures, except for those in the direction of output 2 and  $\tau_E$ . Posterior mean environmental efficiency [see (28)] is clearly negatively correlated with the other efficiency measures, and in particular, has a correlation of  $-0.51$  with the technical efficiency for the good outputs,  $\tau_T$ . This result is in contrast with the small positive correlation found in Fernández *et al.* (2002). Note, however, that in the latter paper environmental and technical efficiency are modeled separately, whereas here they reflect different ways of looking at the same distance with respect to the (single) frontier.

The relationship between  $\exp(-\gamma)$ <sup>7</sup> and the various input-oriented efficiency measures is further illustrated in Figures 9-12, where it is clear that the relationship bears some resemblance to the powers of  $\exp(-\gamma)$  that result from the simpler Cobb-Douglas frontier [see (18) and (19)].

Table 8 is the counterpart of Table 4, where we have chosen three quartile firms according to  $\exp(-\gamma)$ , and consider the posterior probabilities that a higher ranked firm is more efficient than a lower ranked one.

Firm A/Firm B	$\exp(-\gamma)$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{RI}$	$\tau_{1I}$	$\tau_{2I}$	$\tau_{3I}$	$\tau_T$	$\tau_E$
3rd Quart./Median	0.99	0.99	0.97	0.99	0.78	0.83	0.85	0.33	0.86	0.99	1.00
3rd Quart./1st Quart.	1.00	1.00	0.01	1.00	0.00	1.00	0.94	0.81	0.91	1.00	1.00
Median/1st Quart.	0.99	1.00	0.00	1.00	0.00	1.00	0.77	0.99	0.39	0.99	0.00

<sup>7</sup>Note that here we do not use  $\tau_{RO}$  since the aggregator function is different from the one in (1), which underlies the logic of (9).

Table 8 paints the same broad picture as Table 7, in that the efficiency measures are roughly in accordance, except for output 2-oriented and environmental efficiency. In order to get an idea of typical regions for the various efficiencies, Table 9 presents the averages and standard deviations (computed over all observations) of the posterior mean efficiencies. Clearly, the inequality constraints mentioned in Section 3 can be recovered in the means.

	$\exp(-\gamma)$	$\tau_{1O}$	$\tau_{2O}$	$\tau_{1S}$	$\tau_{2S}$	$\tau_{RI}$	$\tau_{1I}$	$\tau_{2I}$	$\tau_{3I}$	$\tau_T$	$\tau_E$
Average	0.68	0.65	0.18	0.60	0.07	0.63	0.11	0.41	0.29	0.67	0.15
St.Dev.	0.13	0.14	0.14	0.14	0.06	0.17	0.11	0.19	0.20	0.13	0.12

Note the large difference between the output-oriented efficiency measures corresponding to the two good outputs. Clearly, firms tend to be much more efficient in the direction of milk production (output 1) than in that of non-milk production. Given that the primary goal of these farms is milk production, their infrastructure will be mostly geared towards that aim, and this difference is not too surprising. The average of the posterior mean efficiencies in Table 9 would correspond to a firm that is much closer to the horizontal (“milk”) axis than the vertical (“non-milk”) axis in Figure 1b, and produces 65% ( $\tau_{1O}$ ) of the milk output that it could produce, given the current level of non-milk production. If it would also relinquish its non-milk production, it is at 60% ( $\tau_{1S}$ ) of its maximum milk output. However, if this firm would consider giving up its milk production, it could multiply its production of non-milk output more than tenfold (see  $\tau_{2S}$ )!

Radial input-oriented efficiency is in the same region as  $\exp(-\gamma)$ , and the input that seems least efficiently used is the first one, namely family labor (which presumably is often readily available).

Of particular interest are also the technical and environmental efficiencies, and from Table 9 we clearly see that firms tend to do better in producing goods than in avoiding the production of the bad (nitrogen surplus). A more insightful picture of technical and environmental efficiencies can be obtained by a predictive concept. The out-of-sample predictive efficiency distributions capture the inference on efficiencies for an unobserved firm belonging to the industry. This is obtained by integrating out the efficiency distributions with the posterior. Given the definition of  $\tau_T$  and  $\tau_E$ , we need to integrate the distribution of  $z$  (and thus  $\gamma$ ) in (A.2) with the posterior distribution of  $\phi$  and the distribution of  $\eta$  in (7) with the posterior of  $s$ , as well as use the posterior of  $\psi$  and  $\alpha$  to define  $q(x)$  in (2)<sup>8</sup> and  $\eta$  in (6). Given posterior draws from  $(\phi, s, \psi, \alpha)$ , we

---

<sup>8</sup>Here  $q(x)$  is simply equal to  $(1 + \psi_0)$  as we are in the separable case. In general,  $q(x)$  does depend on inputs, and then we could *e.g.* present the predictive for a farm with mean or median values for the inputs.

can thus construct draws for technical and environmental efficiencies. This results in the predictive efficiencies graphed in Figures 13 and 14, along with the efficiencies of the three quartile observations chosen for Table 8. The latter illustrate that knowledge of the specific firm-year greatly reduces the uncertainty in efficiency inference. This, once again, illustrates the possibility of ranking firms according to efficiencies in a meaningful way (although we should not attach too much importance to small differences in rank). Predictive means and standard deviations (in parentheses) of  $\tau_T$  and  $\tau_E$  are 0.701 (0.220) and 0.244 (0.221), respectively. If we compare these predictive distributions with those presented in Fernández *et al.* (2002), we obtain here a slightly higher technical efficiency and a somewhat lower environmental efficiency. The overall picture, though, is similar, in that uncertainty about efficiencies of unobserved firms is substantial and environmental efficiency tends to be considerably lower than technical efficiency.

It is worth noting that this paper focuses on multiple-output production. If there is only a single output, then all of our output-oriented measures are equivalent to one another. However, the input-oriented measures are still different from one another. Hence, the issues and findings in this application for the input-oriented measures are of relevance for the single-output case. Atkinson and Cornwell (1994) provide a detailed discussion of output and input-oriented measures in the single equation case.

## 6 Conclusions

In this paper, we have derived and discussed various ways of defining efficiency with multiple-output production, where we allow for nonseparability of outputs and inputs. In addition to standard radial measures we discuss the properties of individual output-oriented measures. We also considered analogous input-oriented efficiency definitions. Furthermore, we developed an efficiency measure which incorporates the losses due to jointness in production, and which we refer to as “efficiency shares”. Finally, we defined technical and environmental efficiencies in cases where some of the outputs are undesirable by-products. With so many alternative ways of defining efficiency, the question arises as to which definition to adopt. This paper provides two types of responses to this. Firstly, in many applications the choice of efficiency measure is not arbitrary. Each measure has a particular interpretation and it is crucial to use the measure that corresponds to the particular empirical question being raised. For instance, the researcher may have reason to focus on one particular output or input and, hence, use  $\tau_{jO}$  or  $\tau_{jI}$ . Alternatively, the researcher may be interested in what would happen if only one output were produced, in which case the concept of efficiency shares may be the relevant one. Secondly, considering more

than one efficiency measure allows us to obtain a much clearer picture of the data, through the decomposition of the distance to the frontier ( $\gamma$ ) in various complementary and often insightful ways.

We have illustrated the issues above in the context of two substantive applications, where we also demonstrated that the types of things a policymaker might want to report (*e.g.* an efficiency ranking or the identification of some firms as laggards or exemplars) do seem to make practical sense, once we have agreed on a particular choice of efficiency definition.

Nevertheless, our results do highlight that caution should be taken when carrying out an efficiency analysis. We would recommend that practitioners report posterior standard deviations (or confidence intervals) of all firm-specific efficiency estimates. Furthermore, if theoretical considerations do not provide clear guidance as to which efficiency definition should be used, researchers should present results for several different choices.

## Acknowledgements

We acknowledge funding by the UK Economic and Social Research Council under grant number R34515. We are grateful to two referees and a guest co-editor for helpful and thought-provoking comments that have led to improvements in the paper. We would also like to thank Robert Adams, Allen Berger and Robin Sickles for providing the banking data used in this paper, and Stijn Reinhard and the Dutch Agricultural Economics Research Institute (LEI-DLO) for providing the agricultural data.

## Appendix: Bayesian Analysis

### The Prior Distribution

In order to conduct Bayesian inference, we need to complement the likelihood function (as described in Section 2) with a prior distribution on the parameters  $(\beta, z, \sigma, \alpha, q, s)$ . We shall choose a proper prior distribution with the following structure

$$p(\beta, z, \sigma, \alpha, q, s) = p(\beta)p(z)p(\sigma)p(\alpha)p(q)p(s). \quad (\text{A.1})$$

It is worthwhile briefly noting that use of improper priors in stochastic frontier models can cause problems (*i.e.* inability to calculate meaningful Bayes factors or even the lack of existence of the posterior itself, see Fernández *et al.*, 1997, for details). In this paper, we choose values for hyperparameters which imply a relatively noninformative but proper prior.

**Prior for  $z$ :**

The distribution of  $z$  determines the distribution of the inefficiency vector  $\gamma = Dz$ . Since we are assuming an individual effects structure for the inefficiencies,  $z \in \mathfrak{R}_+^N$ . We will consider a hierarchical structure which adds an extra parameter  $\phi \in \mathfrak{R}_+$ . Given  $\phi$ , we take the following product of Exponential distributions for the elements of  $z$ :

$$p(z|\phi) = \prod_{l=1}^N f_G(z_l|1, \phi), \quad (\text{A.2})$$

where  $f_G(z_l|a, b)$  denotes the p.d.f. of a Gamma distribution with mean  $a/b$  and variance  $a/b^2$ . The prior on  $\phi$  is taken to be

$$p(\phi) = f_G(\phi|e_0, g_0), \quad (\text{A.3})$$

with positive prior hyperparameters  $e_0$  and  $g_0$ . We choose  $e_0 = 1$  and  $g_0 = -\log(0.80)$ , values which imply a relatively flat prior on the individual efficiencies with prior median efficiency set at 0.80, as discussed in van den Broeck *et al.* (1994).

**Prior for  $\beta$ :**

The prior assumed on the frontier parameters has p.d.f.

$$p(\beta) \propto f_N^k(\beta|b_0, H_0^{-1})I_{\mathcal{B}}(\beta),$$

*i.e.* a  $k$ -dimensional Normal distribution with mean  $b_0$  and covariance matrix  $H_0^{-1}$ , truncated to the regularity region  $\mathcal{B}$ . In our first application, we assume a Cobb-Douglas form for the frontier. Hence,  $I_{\mathcal{B}}(\beta)$  corresponds to simply restricting the elements of  $\beta$  (except the intercept) to be non-negative. In our second application, we use a translog frontier and we impose only local regularity (*i.e.* we impose regularity at each data point). Hence,  $I_{\mathcal{B}}(\beta)$  corresponds to the region where input elasticities are non-negative at every data point [see the discussion after equation (22)]. We set  $b_0 = 0_k$  and  $H_0 = 10^{-4} \times I_k$ . The  $k - 1$  last elements of  $\beta$  are likely to be smaller than 1. Due to the logarithmic transformation, the intercept will also certainly be much less than one prior standard deviation away from the mean. Hence, our prior is quite noninformative.

**Prior for  $\sigma$ :**

We define the prior distribution on the scale  $\sigma$ , through a Gamma distribution on the precision  $h = \sigma^{-2}$ :

$$p(h) = f_G(h|n_0/2, a_0/2). \quad (\text{A.5})$$

In our applications, we set  $n_0/2 = 1$  (which leads to an Exponential prior for  $h$ ) and  $a_0/2 = 10^{-6}$ . These values imply large prior uncertainty.

**Prior for  $\alpha$ :**

Since the components of  $\alpha$  are all in the interval  $(0, 1)$  and sum up to one, an obvious choice of a prior distribution is a Dirichlet with p.d.f.

$$p(\alpha) = f_D^{p-1}(\alpha|a), \quad (\text{A.6})$$

where the hyperparameter  $a = (a_1, \dots, a_p)' \in \mathfrak{R}_+^p$ . We use  $a = \iota_p$ , which makes the prior Uniform over the  $(p - 1)$ -dimensional unit simplex.

**Prior for  $\psi$ :**

For the parameterization of  $q$ , we use independent Gamma priors

$$p(\psi) \propto \prod_{l=0}^m f_G(\psi_l|d_1, d_2). \quad (\text{A.7})$$

In our empirical illustrations, we choose  $d_1 = d_2 = 1$ .

**Prior for  $s$ :**

We assume  $p$  independent Gamma distributions for the components of  $s$ :

$$p(s) = \prod_{j=1}^p p(s_j) = \prod_{j=1}^p f_G(s_j|b_j, c_j). \quad (\text{A.8})$$

In our applications, we take  $b_j = 1$  and  $c_j = 10^{-6}$  for all  $j$ .

**The MCMC Sampler**

The likelihood function derived from (5) and (7) is given in expression (2.7) of FKS. In combination with the prior distribution described in the first part of this Appendix, this defines a Bayesian model. In order to conduct posterior inference with this model, we shall use an MCMC sampler to generate drawings from it. The sampler is run over the inefficiencies in  $z$  as well as all parameters, including the parameter  $\phi$  introduced in the prior for  $z$ . We shall briefly describe the sampler here and refer to FKS for more details in a very similar context. If not otherwise indicated, indices run from  $i = 1, \dots, N$ ,  $t = 1, \dots, T$  and  $j = 1, \dots, p$ . The full conditionals used in the sampler are as follows:

**Conditional posterior distribution of  $z$ :**

Choosing  $D$  in (5) to ensure constant efficiencies over time, we obtain

$$p(z|Y, \phi, \beta, \sigma, \alpha, \psi, s) \propto f_N^N(z|m, R)I_{\mathfrak{R}_+^N}(z), \quad (\text{A.9})$$

*i.e.* a truncated  $N$ -variate Normal distribution, with mean

$$m = (D'D)^{-1}\{D'(V\beta - \delta) - \sigma^2\phi\iota_N\}, \quad (\text{A.10})$$

and covariance matrix

$$R = \sigma^2(D'D)^{-1}. \quad (\text{A.11})$$

Note that in our case  $D'D = \text{Diag}(T_i)$  is diagonal, which greatly facilitates drawing from (A.9).

**Conditional posterior distribution of  $\phi$ :**

The conditional distribution of  $\phi$  is given by the p.d.f.

$$p(\phi|Y, z, \beta, \sigma, \alpha, \psi, s) = f_G(\phi|e_0 + N, g_0 + z'\iota_N). \quad (\text{A.12})$$

**Conditional posterior distribution of  $\beta$ :**

For  $\beta$  we get the following truncated Normal distribution

$$p(\beta|Y, z, \phi, \sigma, \alpha, \psi, s) \propto f_N^k(\beta|b_*, H_*^{-1})I_{\mathcal{B}}(\beta), \quad (\text{A.13})$$

where

$$H_* = H_0 + \sigma^{-2}V'V, \quad (\text{A.14})$$

and

$$b_* = H_*^{-1}\{H_0b_0 + \sigma^{-2}V'(\delta + Dz)\}. \quad (\text{A.15})$$

**Conditional posterior distribution of  $\sigma$ :**

The conditional posterior distribution expressed in terms of  $\sigma^{-2}$  takes the simple form

$$p(\sigma^{-2}|Y, z, \phi, \beta, \alpha, \psi, s) = f_G\left(\sigma^{-2}\left|\frac{n_0 + NT}{2}, \frac{a_0 + (\delta - V\beta + Dz)'(\delta - V\beta + Dz)}{2}\right.\right). \quad (\text{A.16})$$

**Conditional posterior distribution of  $\alpha$ :**

The  $p$ -dimensional vector  $\alpha$  has a  $(p - 1)$ -dimensional distribution, where each of the components are in between zero and one and sum to one. We obtain the following p.d.f. for this conditional distribution:

$$p(\alpha|Y, z, \phi, \beta, \sigma, q, s) \propto \prod_j \alpha_j^{a_j - 1 + s_j \sum_{i,t} q(x_{(i,t)})} \prod_{i,t} \left( \sum_j \alpha_j^{q(x_{(i,t)})} y_{(i,t,j)} \right)^{-\sum_j s_j} \exp\left\{-\frac{1}{2\sigma^2}(\delta - V\beta + Dz)'(\delta - V\beta + Dz)\right\}, \quad (\text{A.17})$$

with  $\delta$ , defined through (1) and (3), depending on  $\alpha$ .

### Conditional posterior distribution of $\psi$ :

For each of the components  $\psi_l, l = 0, \dots, m$  we have the posterior

$$p(\psi_l | Y, z, \phi, \beta, \sigma, \alpha, \{\psi_h, h \neq l\}, s) \propto \psi_l^{d_1-1} \exp(-d_2 \psi_l) (1 + \psi_l)^{(p-1) \sum_{i,t} x_{(i,t,l)}} \exp \left\{ -\frac{1}{2\sigma^2} (\delta - V\beta + Dz)' (\delta - V\beta + Dz) \right\} \prod_{i,t,j} \eta_{(i,t,j)}^{s_j}, \quad (\text{A.18})$$

where we have defined  $x_{(i,t,0)} = 1$ . Note that  $\delta$  depends on  $\psi$  through (1)-(3) and  $\eta_{i,t,j}$  depends on  $\psi$  through (6) and (2).

### Conditional posterior distribution of $s$ :

Finally, the p.d.f. of the conditional posterior distribution of  $s_j, j = 1, \dots, p$  is:

$$p(s_j | Y, z, \phi, \beta, \sigma, \alpha, q, \{s_h : h \neq j\}) \propto \left\{ \frac{\Gamma(\sum_l s_l)}{\Gamma(s_j)} \right\}^{NT} s_j^{b_j-1} \exp \left[ -s_j \left\{ c_j - \sum_{i,t} \log \eta_{(i,t,j)} \right\} \right]. \quad (\text{A.19})$$

A simple random walk Metropolis-Hastings algorithm (see *e.g.* Chib and Greenberg, 1995, for an intuitive explanation) with a Normal candidate generator is used to implement the steps for  $\alpha, \psi$  and  $s$  in the sampler. Direct draws from the other conditionals can easily be conducted. Preliminary runs are used to calibrate the variances/covariances for our Normal candidate generating densities used in the Metropolis-Hastings algorithms. The algorithm converges quite rapidly as shown in FKS.

## References

- Adams, R., Berger, A., Sickles, R., 1999, Semiparametric approaches to stochastic panel frontiers with applications in the banking industry, *Journal of Business and Economic Statistics* 17, 349-358.
- Atkinson, S., Cornwell, C., 1994, Estimation of output and input technical efficiency using a flexible functional form and panel data, *International Economic Review*, 35, 231-244.
- Atkinson, S., Cornwell, C., Honerkamp, O., 2003, Measuring and decomposing productivity change: Stochastic distance function estimation versus data envelopment analysis, *Journal of Business and Economic Statistics* 21, 284-294.
- Ball, E., Färe, R., Grosskopf, S., Nehring, R., 2001, Productivity of the U.S. agricultural sector: The Case of undesirable outputs, in: C. Hulten, E. Dean and M. Harper, eds., *New developments in productivity analysis* (Chicago, University of Chicago Press: Chicago) 541-586.
- Barla, P., Perelman, S., 1989, Technical efficiency in airlines under regulated and deregulated environments, *Annals of Public and Cooperative Economics* 6, 61-80.

- Berger, A., 1993, "Distribution-free" estimates of efficiency in the U.S. banking industry and tests of the standard distributional assumptions, *Journal of Productivity Analysis* 4, 261-292.
- Berger, A., Humphrey, D., 1997, Efficiency of financial institutions: International survey and directions for future research, *European Journal of Operational Research* 98, 175-212.
- van den Broeck, J., Koop, G., Osiewalski J., Steel, M.F.J., 1994, Stochastic frontier models: A Bayesian perspective, *Journal of Econometrics* 61, 273-303.
- Chambers, R., Chung, Y., Färe, R., 1996, Benefit and distance functions, *Journal of Economic Theory* 70, 407-419.
- Chib, S., Greenberg, E., 1995, Understanding the Metropolis-Hastings algorithm, *The American Statistician* 49, 327-335.
- Coelli, T., Perelman, S., 2000, Technical efficiency of European railways: A Distance function approach, *Applied Economics* 32, 1967-1976.
- Färe, R., Grosskopf, S., Kirkley, J., 2000, Multi-output capacity measures and their relevance for productivity, *Bulletin of Economic Research* 52, 1-12.
- Färe, R., Grosskopf, S., Noh, D.-W., Weber, W., 2002, Characteristics of a Polluting Technology: Theory and Practice, manuscript, Oregon State University, Department of Economics.
- Färe, R., Grosskopf, S., Zaim, O., 1999, An index number approach to measuring environmental performance: An environmental Kuznets curve for the OECD countries, manuscript, Oregon State University, Department of Economics.
- Färe, R., Primont, D., 1995, *Multi-output production and duality: Theory and applications* (Kluwer, Boston).
- Fernández, C., Koop, G., Steel, M.F.J., 2000, A Bayesian analysis of multiple-output production frontiers, *Journal of Econometrics* 98, 47-49.
- Fernández, C., Koop, G., Steel, M.F.J., 2002, Multiple-output production with undesirable outputs: An application to nitrogen surplus in agriculture, *Journal of the American Statistical Association* 97, 432-442.
- Fernández, C., Osiewalski, J., Steel, M.F.J., 1997, On the use of panel data in stochastic frontier models with improper priors, *Journal of Econometrics* 79, 169-193.
- Horrace, W., Schmidt, P. 1996, Confidence statements for efficiency estimates from stochastic frontier models, *Journal of Productivity Analysis* 7, 257-282.
- Hunt-McCool, J., Koh, S.C., Francis, B.B., 1996, Testing for deliberate underpricing in the IPO premarket: A stochastic frontier approach, *Review of Financial Studies* 9, 1251-1269.

- Kim, Y., Schmidt, P., 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.
- Koop, G., 1998, Carbon Dioxide emissions and economic growth: A structural approach, *Journal of Applied Statistics* 25, 489-515.
- Koop, G., Osiewalski, J., Steel, M.F.J., 1994, Bayesian efficiency analysis with a flexible cost function, *Journal of Business and Economic Statistics* 12, 93-106.
- Koop, G., Osiewalski, J., Steel, M.F.J., 1997, Bayesian efficiency analysis through individual effects: Hospital cost frontiers, *Journal of Econometrics* 76, 77-105.
- Kumbhakar, S., 1987, The specification of technical and allocative inefficiency of multi-product firms in stochastic production and profit frontiers, *Journal of Quantitative Economics*, 3, 213-223.
- Kumbhakar, S., Ghosh, S., McGuckin, J.T., 1991, A generalized production frontier approach for estimating determinants of inefficiency in U.S. dairy farms, *Journal of Business and Economic Statistics* 9, 279-286.
- Newton, M., Raftery, A., 1994, Approximate Bayesian inference by the weighted likelihood bootstrap," *Journal of the Royal Statistical Society, B*, 56, 3-48 (with discussion).
- Perelman, S., Pestieau, P., 1994, A comparative performance study of postal services: A productive efficiency approach, *Annals d'Economie et de Statistique* 33, 187-202.
- Pitt, M., Lee, L.-F., 1981, The measurement and sources of technical inefficiency in the Indonesian weaving industry, *Journal of Development Economics* 9, 43-64.
- Poirier, D., 1995, *Intermediate statistics and econometrics* (MIT Press: Cambridge).
- Powell, A., Gruen, F., 1968, The constant elasticity of transformation production frontier and linear supply system, *International Economic Review*, 9, 315-328.
- Reinhard, S., 1999, *Econometric analysis of economic and environmental efficiency of Dutch dairy farms* (PhD dissertation, Wageningen University).
- Reinhard, S., Lovell, C.A.K., Thijssen, G., 1999, Econometric application of technical and environmental efficiency: An application to Dutch dairy farms, *American Journal of Agricultural Economics* 81, 44-60.
- Shephard, R.W., 1970, *Theory of cost and production functions* (Princeton University Press, Princeton).

Tyteca, D., 1996, On the measurement of the environmental performance of firms: A literature review and a productive efficiency perspective, *Journal of Environmental Management* 46, 281-308.

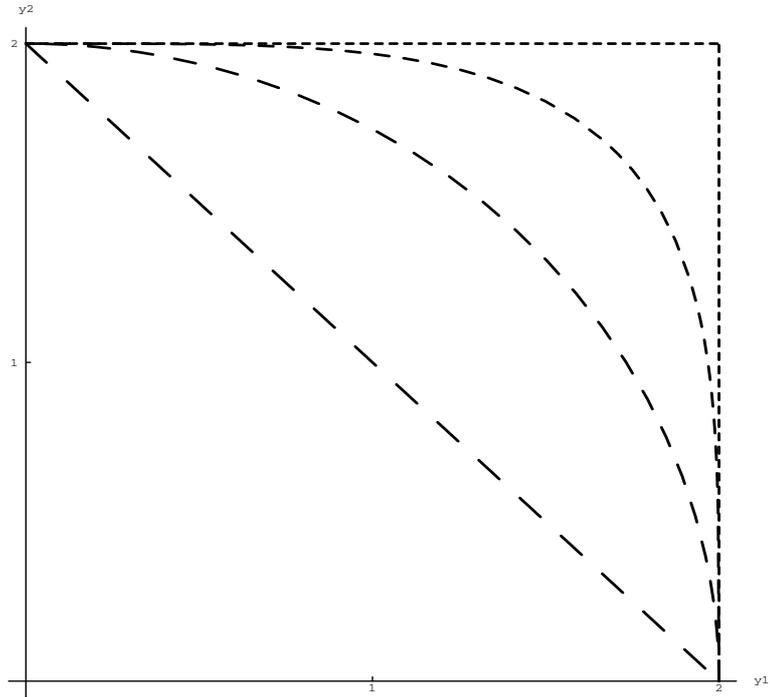


Figure 1a: Production equivalence surface with  $p = 2$  outputs and various values of  $q(x_{(i,t)})$ .  $q(x_{(i,t)}) = 1$  for the longest dashes and as the dashing becomes finer, we use  $q(x_{(i,t)}) = 2$ ,  $q(x_{(i,t)}) = 4$  and, finally,  $q(x_{(i,t)}) = \infty$  for the shortest dashing. We have taken  $\alpha_1 = \alpha_2 = 1/2$  and  $g(y_{(i,t)}, x_{(i,t)}) = 1$ .

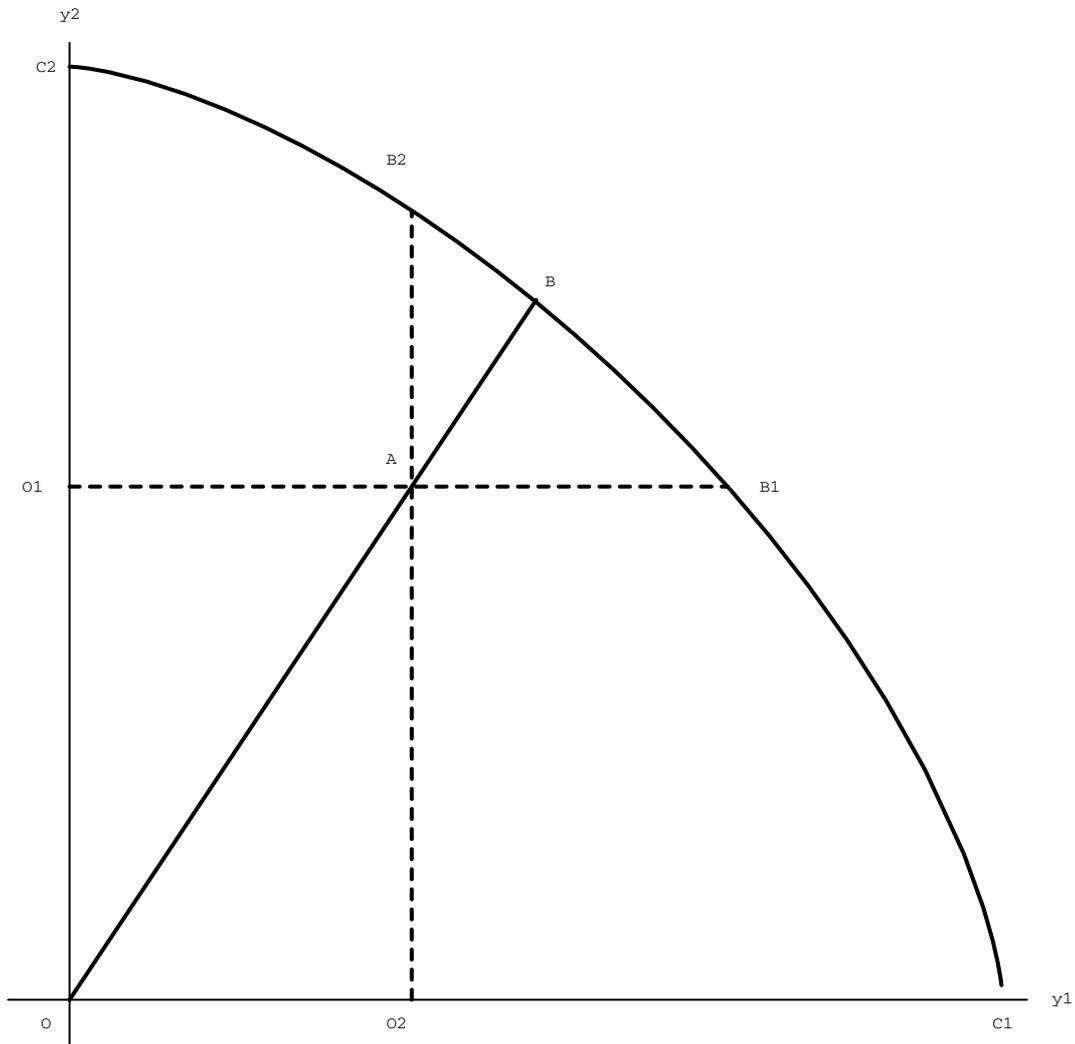


Figure 1b: Production equivalence surface with  $p = 2$  good outputs and  $q = 1.5$ . A firm producing  $A$  will have radial output-oriented efficiency equal to  $|OA|/|OB|$ , whereas individual output-oriented efficiencies are given by  $\tau_{1O} = |O_1A|/|O_1B_1|$  and  $\tau_{2O} = |O_2A|/|O_2B_2|$ . Finally, efficiency shares are equal to  $\tau_{1S} = |OO_2|/|OC_1|$  for output  $y_1$  and to  $\tau_{2S} = |OO_1|/|OC_2|$  for the second output.

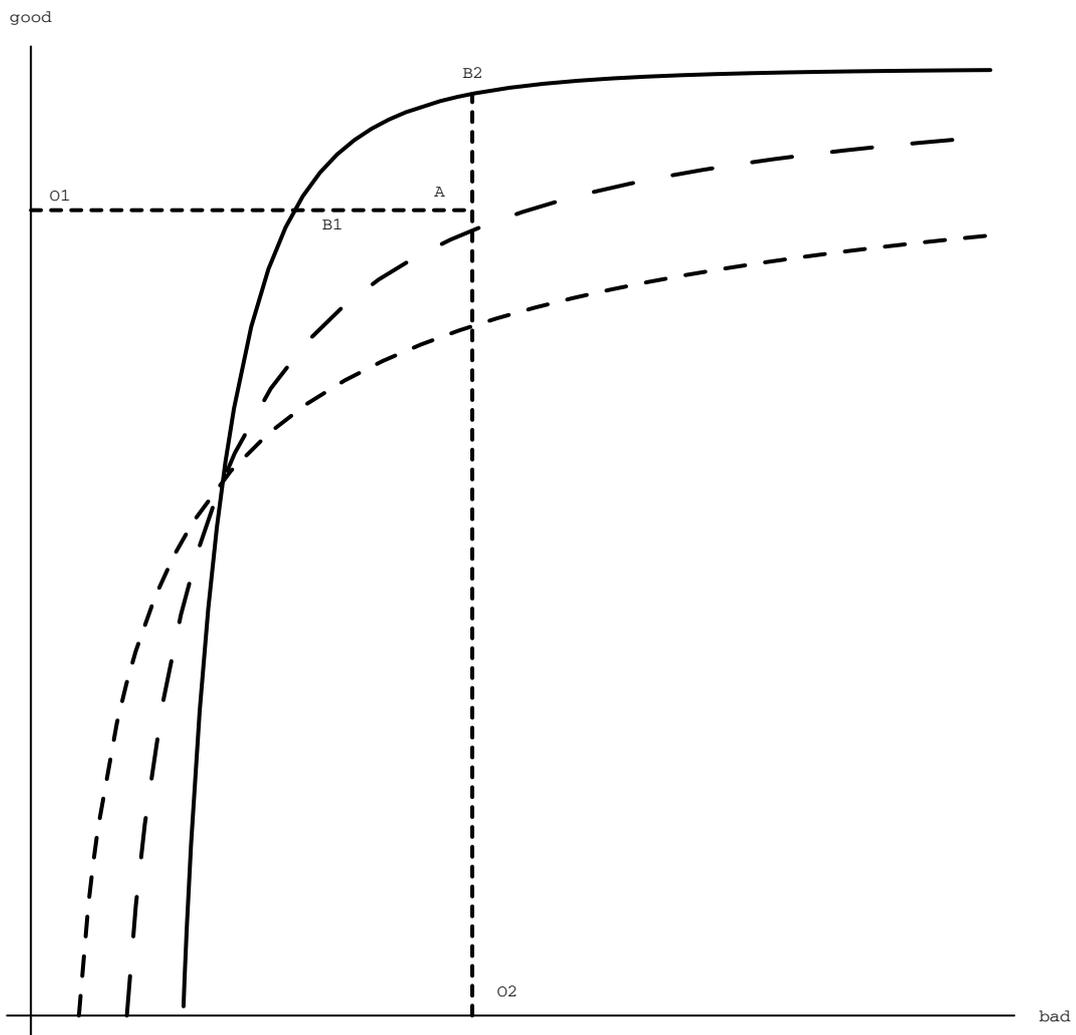


Figure 2: Production equivalence surface with one good and one bad output for  $q = 1.1$ . The drawn line corresponds to  $r = 3$ , the long dashes to  $r = 1$  and the short dashes to  $r = 0.5$ . Production at point  $A$  corresponds to technical efficiency  $|O_2A|/|O_2B_2|$ , whereas the environmental efficiency is given by  $\tau_E = |O_1B_1|/|O_1A|$ .

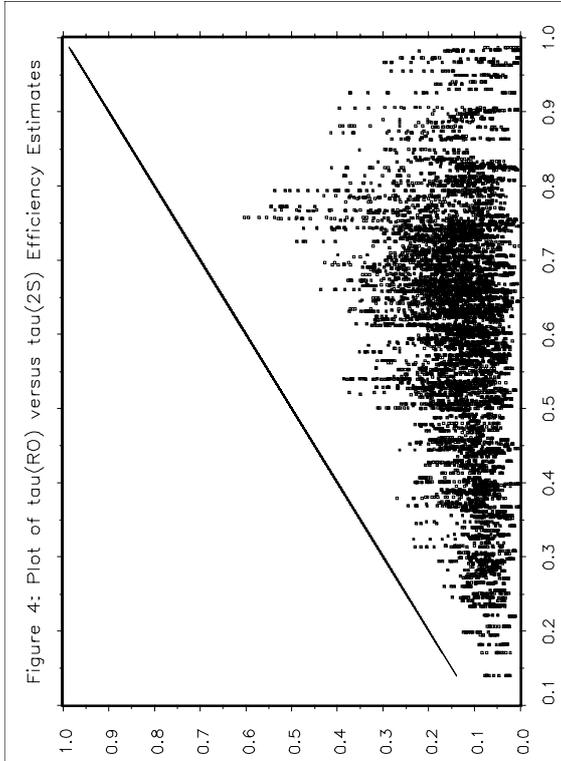


Figure 4: Plot of  $\tau(\text{RO})$  versus  $\tau(2\text{S})$  Efficiency Estimates

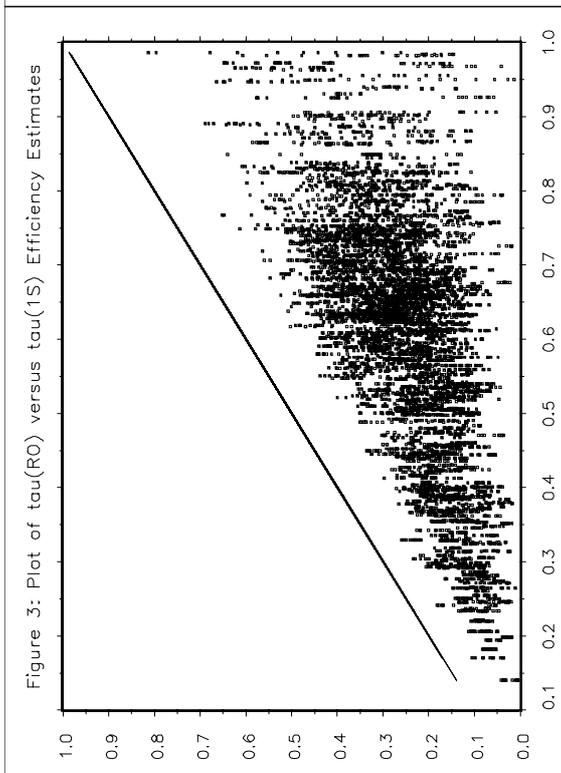


Figure 3: Plot of  $\tau(\text{RO})$  versus  $\tau(1\text{S})$  Efficiency Estimates

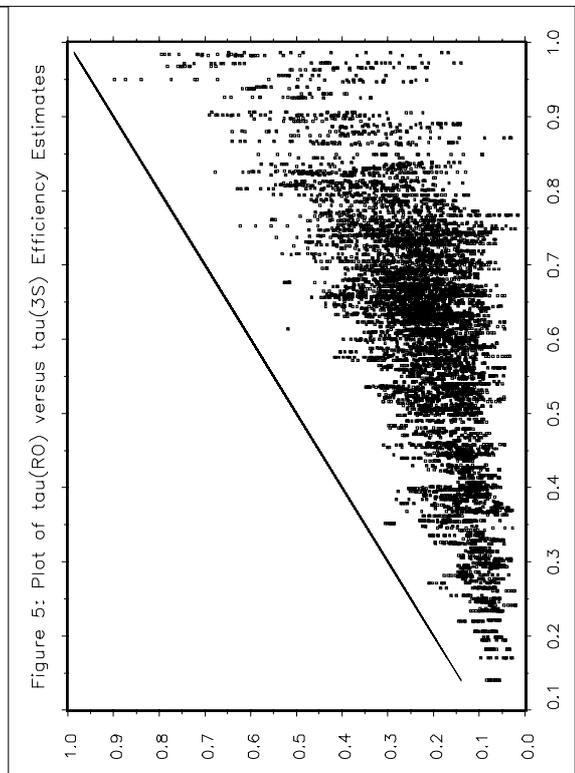
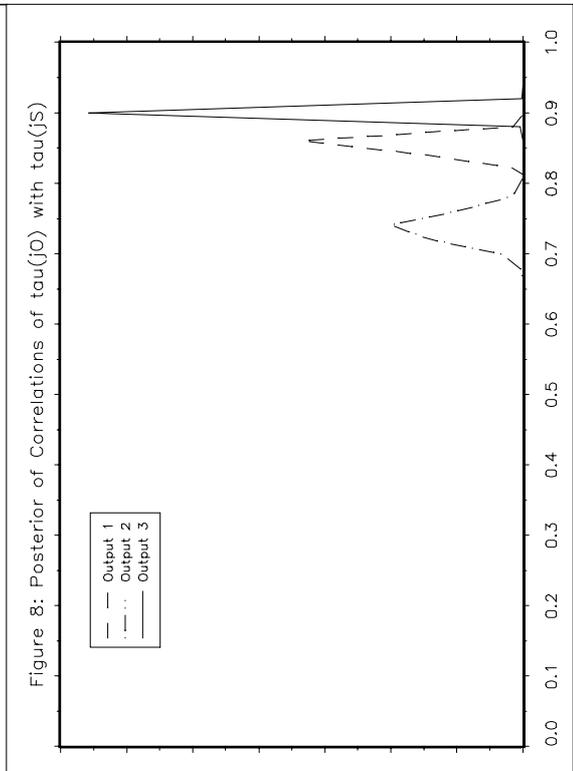
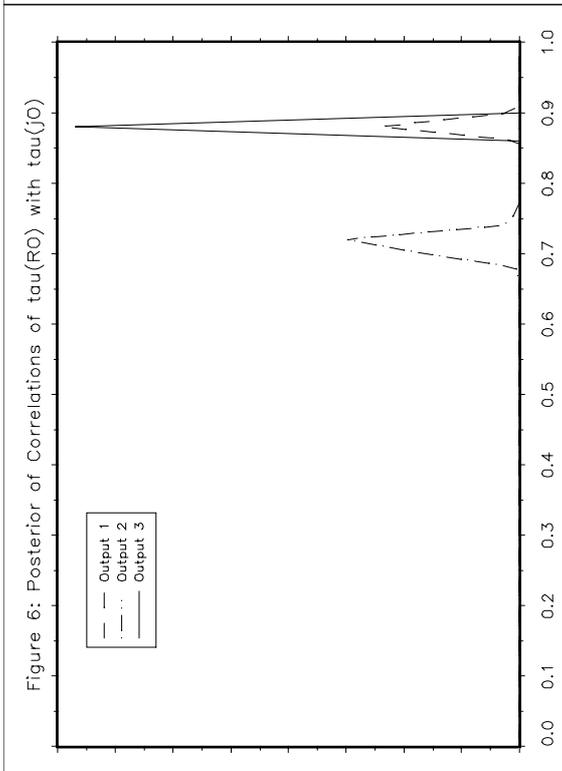
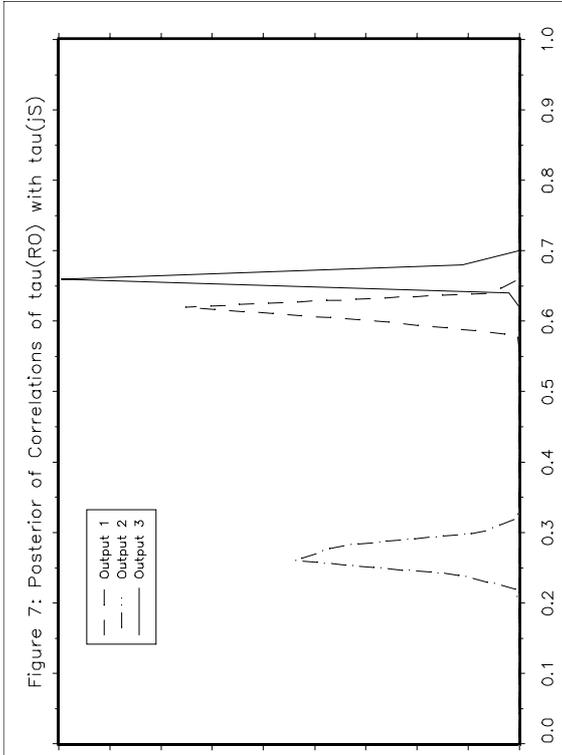
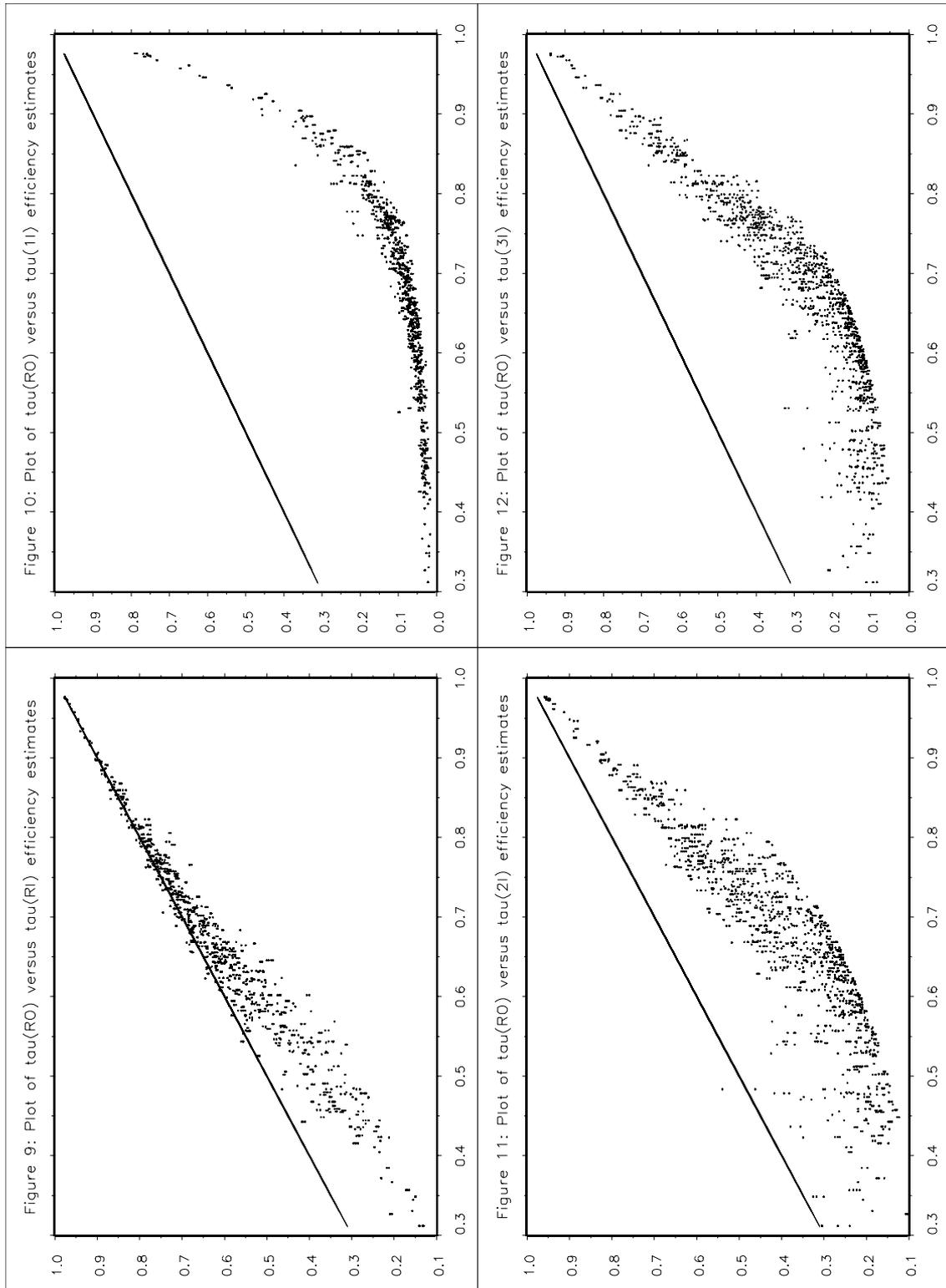


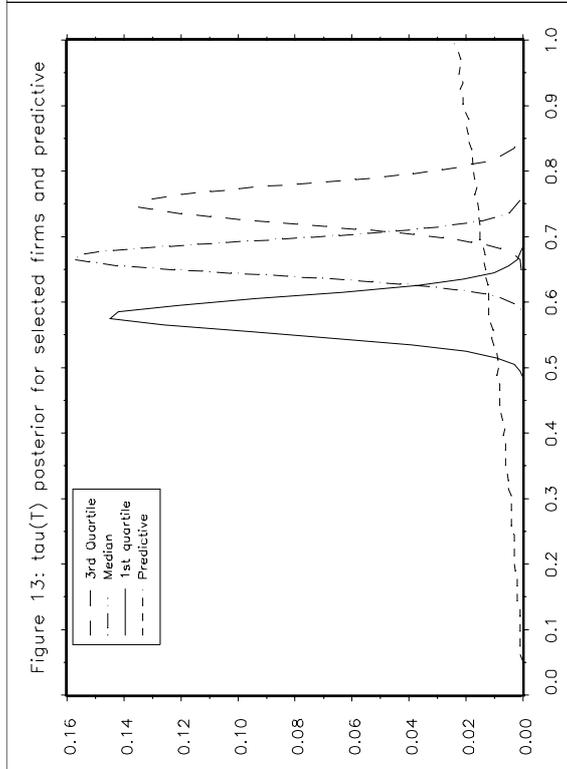
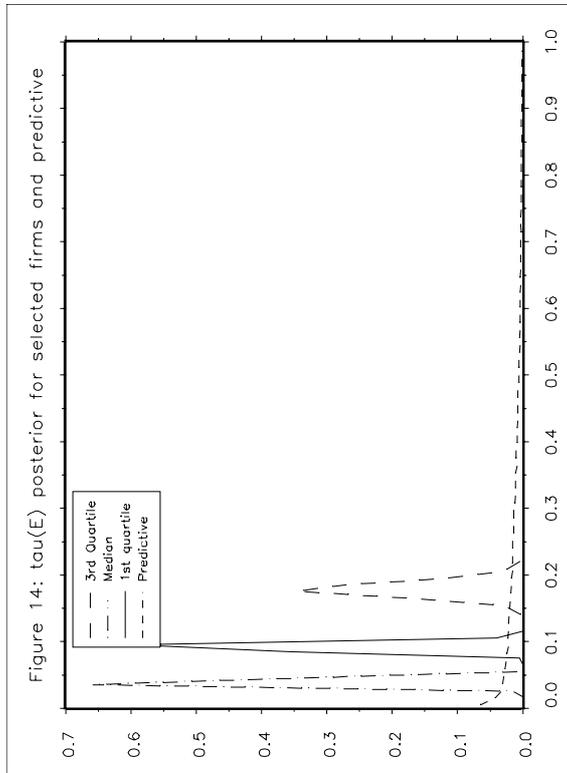
Figure 5: Plot of  $\tau(\text{RO})$  versus  $\tau(3\text{S})$  Efficiency Estimates



Figures 6-8: Posterior densities of correlations between output-oriented efficiency measures for the banking data



Figures 9-12: Posterior means of radial output-oriented efficiency  $\exp(-\gamma)$  versus input-oriented efficiency measures for the farm data



Figures 13 and 14: Posterior and predictive densities of technical and environmental efficiencies for the farm data