Estimating the Characteristics of Polluting Technologies*

C.J. O'Donnell
University of Queensland
email: c.odonnell@economics.uq.edu.au

January 21, 2007

Abstract

Polluting technologies can be represented using output distance functions. A common approach to estimating such functions is to factor out one of the outputs and estimate the resulting equation using well-known stochastic frontier estimation methods, including maximum likelihood. A problem with this approach is that the outputs that are not factored out may be correlated with the error term, leading to biased and inconsistent estimates. This paper addresses the problem in a Bayesian framework. The methodology is applied to data on U.S. electric utilities. Results include estimates of technical inefficiencies and the shadow price of a pollutant.

1 Introduction

The economically-relevant characteristics of a multi-input multi-output production technology can be represented using a Shephard output distance function. This function gives the inverse of the largest factor by which a firm can radially expand its output vector while holding its input vector fixed. The econometric approach to estimating Shephard output distance functions typically involves factoring out one of the outputs and estimating the resulting equation using conventional stochastic frontier estimation methods, such as corrected ordinary least squares (COLS) or maximum likelihood. Examples include Reinhard and Thijssen (1998), Coelli and Perelman (1999) and O'Donnell and Coelli (2005).

A directional output distance function is a slightly different type of distance function that can be used to represent technologies that are non-separable in desirable and undesirable outputs. Examples of undesirable outputs include

---

workplace injuries and agricultural and industrial pollutants. The directional distance function is useful in these production contexts because it allows for non-radial expansions of the output vector. Thus, it can register a performance improvement when firms increase some outputs (the desirable ones) and decrease others (the pollutants). Recently, Fare et. al. (2005) used a directional output distance function to measure the performance of firms producing sulphur dioxide (SO_2) as a byproduct of electricity generation. Estimation was effected by factoring out the SO_2 output and estimating the resulting parametric frontier model using COLS.

Estimating output distance functions in a parametric frontier framework is far from straightforward. A problem that concerns many researchers is that the outputs that are not factored out may be correlated with the composite error term (Atkinson, Fare and Primont, 1998; Atkinson and Primont, 1998). The problem is sometimes referred to as the 'endogeneity problem' (e.g., Robias and Arias, 2005). One solution is to estimate the frontier using the generalized method of moments (GMM) (e.g., Atkinson, Cornwell and Honerkamp, 2003). An advantage of this approach is that it obviates the need for distributional assumptions on the error term. Unfortunately, GMM estimates are often sensitive to the choice of instruments, and the finite sample properties of the estimator are unknown.

In this paper, we address the endogeneity problem in a way that does not involve the use of instruments. Our alternative approach involves the imposition of inequality constraints on the parameters of an approximating functional form. Importantly, the approach can be implemented in a Bayesian framework. One advantage of the Bayesian approach is that it enables us draw exact finite sample inferences concerning nonlinear functions of the unknown parameters. As an illustration, we use the approach to draw inferences concerning the technical inefficiencies of U.S. electric utilities. We also draw exact finite sample inferences concerning the shadow price of the pollutant.

The structure of the paper is as follows. Section 2 formally introduces the directional output distance function and its properties. These properties include representation and monotonicity. Section 3 shows how the representation property can be used to derive an empirical model that can be estimated using either sampling theory (i.e., GMM) or Bayesian methods. However, to avoid problems associated with choosing instruments, and to facilitate the imposition of monotonicity constraints, this paper adopts the Bayesian approach. Section 4 derives the joint posterior density for the unknown parameters and unobserved inefficiency effects. The conditional posterior densities to be used in a Gibbs sampling algorithm are also derived. Section 5 describes the data and reports the empirical results. The results are compared to nonparametric estimates obtained by Fare et. al. using the same data set. The paper is concluded in Section 6.

2 The Directional Output Distance Function

Let S_t denote the set of all output and input vectors $y_t \in \mathbb{R}_+^N$ and $x_t \in \mathbb{R}_+^M$ that are technically feasible in period $t \in \{1, ..., T\}$. The production technology can be represented by the directional output distance function

$$D(x_t, y_t; g) = \sup\{\beta : (x_t, y_t + \beta g) \in S_t\} \tag{1}$$

where $g = (g_1, ..., g_M)'$ is a direction vector. The m-th output is considered desirable if $g_m > 0$ and undesirable if $g_m < 0$. If the technology satisfies mild regularity conditions then the following properties hold (e.g., Fare et. al., 2005):

1. Representation: $D(x_t, y_t; g) \geq 0$ if and only if $(x_t, y_t) \in S_t$.

2. Translation: $D(x_t, y_t + \lambda g; g) = D(x_t, y_t; g) - \lambda$.

3. g-homogeneity: $D(x_t, y_t; \lambda g) = \lambda^{-1} D(x_t, y_t; g)$ for $\lambda > 0$.

4. No free lunch: $D(0, 0; g) = 0$.

5. Weak disposability: $D(x_t, \theta y_t; g) \geq 0$ for $(x_t, y_t) \in S_t$ and $0 \leq \theta \leq 1$.

6. Monotonicity: $G\nabla_y D(x_t, y_t; g) \leq 0_M$, where $G = \text{diag}(g_1, ..., g_M)$.

The representation property says that the directional distance function is non-negative for all feasible input-output combinations. It also implies the distance function takes the value zero if and only if the firm lies on the boundary of the production possibilities set. Such a firm is said to be efficient in the g-direction, or g-efficient. The translation property says that expanding the output vector by an amount λg has the effect of reducing the directional distance by λ. This property is the additive analogue of the homogeneity property of the Shephard output distance function. The g-homogeneity property says that a radial expansion of the direction vector causes an equiproportionate reduction in the directional distance, while the no free lunch property says that production of non-zero output is impossible without committing some inputs to the production process. The weak disposability property says that any proportional contraction of outputs is possible. Finally, the monotonicity property says that the directional distance function is non-increasing in desirable outputs and non-decreasing in undesirable outputs.

The Shephard output distance function is a limiting special case of the directional output distance function corresponding to $g = y_t$.

3 The Empirical Model

The representation property can be used to write $D(x_t, y_t; g) = u_t$ where $u_t \geq 0$ is a random variable with mean $\mu \equiv E(u_t) > 0$. Let $F(x_t, y_t; g)$ be a function that can provide a local approximation to $D(x_t, y_t; g)$ at the variable means. We follow standard practice in the applied economics literature and treat the

observation-varying error of approximation, $v_t \equiv D(x_t, y_t; g) - F(x_t, y_t; g)$, as an independent symmetric random variable with zero mean. It follows that

$$F(x_t, y_t; g) = u_t - v_t \tag{2}$$

where $\epsilon_t \equiv u_t - v_t$ is a random variable with mean μ. All flexible functional forms are available for use in empirical work. In this paper we use the quadratic approximating function

$$F(x_t, y_t; g) = \delta + a'y_t + b'x_t + y_t'Ay_t + x_t'Bx_t + x_t'Hy_t \tag{3}$$

where δ is an unknown scalar and a, b, A, B and H are unknown matrices. For identification purposes, the matrices A and B are assumed to be symmetric.

There is no theoretical requirement that the function F exhibit the translation, g-homogeneity or no free lunch properties of the unknown function D. However, if F is to provide a first-order approximation to D at the variable means, it must satisfy the monotonicity property at that point. In this paper, where we normalise all input and output variables to have unit means, the parameters of F must satisfy $G\nabla_y F(\iota_N, \iota_M; g) \leq 0_M$, where ι_N denotes a unit column vector of order N. Unconstrained estimation is possible using GMM, where an obvious moment condition is $E\{F(x_t, y_t; g) - \mu\} = 0$. However, no other moment conditions are immediately apparent. Moreover, there is no satisfactory method for imposing the monotonicity constraint. This sparks our interest in an alternative estimation strategy.

Our alternative strategy involves the construction of a nonnegative aggregate output, $q_t \equiv w_t'y_t$. Since the approximating function is quadratic, it is convenient to define the weight vector

$$w_t \equiv Ay_t - \nabla_y F(x_t, y_t; g) = -(a + Ay_t + Hx_t)' \geq 0_M \tag{4}$$

The assumption that $w_t \geq 0_M$ is particularly useful for estimation purposes, not least because it ensures the aggregate output is non-negative for all $y_t \in \mathbb{R}_+^N$. Also germaine to the estimation problem is the fact that equations (2), (3) and (4) combine to form an empirical model that resembles a conventional stochastic frontier model:

$$q_t = \delta + b'x_t + x_t'Bx_t + v_t - u_t. \tag{5}$$

This equation suggests that the unknown parameters and inefficiency effects can be estimated using conventional frontier estimation techniques. Unfortunately, conventional frontier estimation is complicated by the fact that the aggregate output is unobserved. In this paper, we solve this particular problem in a Bayesian framework.

4   Bayesian Estimation

Bayesian estimation involves sampling from the joint posterior probability density function (pdf) of the unknown parameters and unobserved inefficiency effects. In this section we derive the likelihood function and specify the prior pdf.

We also specify the conditional posteriors needed for a Markov Chain Monte Carlo (MCMC) sampling algorithm.

4.1 The Likelihood Function

The T equations represented by (5) can be written in the compact form

$$q = X\beta + v - u \tag{6}$$

where $q = (q_1, ..., q_T)'$; X is a $T \times (1 + 0.5N)(N + 1)$ matrix with t-th row comprising a constant, the elements of x_t, and the distinct elements of $x_t x_t'$; β is a conformable vector containing δ, the elements of b, and the distinct elements of B; and the remaining definitions are obvious. For inference in models of this type, it is common to assume v is an independent normal random vector with pdf[1]

$$p(v \mid h) = f_N\left(v \mid 0_T, h^{-1}I_T\right) \tag{7}$$

where I_T denotes an identity matrix of order T. Thus, the conditional joint density for the unobserved aggregate outputs is

$$p(q \mid \beta, u, h) = f_N\left(q \mid X\beta - u, h^{-1}I_T\right) \tag{8}$$

where, for notational convenience, the conditioning on X has been suppressed. Unfortunately, this T-variate density is not enough to define a sampling density for the $M \times T$ matrix of observed outputs, $Y = (y_1, ..., y_T)$. To define such a sampling density we must introduce $M - 1$ new random variables to drive stochastics in another $M - 1$ dimensions. We do this using an approach employed in a similar context by Fernandez, Koop and Steel (2000) (hereinafter referred to as FKS).

Let w_{mt} and y_{mt} denote the m-th elements of w_t and y_t, and let $\eta_{mt} \equiv w_{mt}y_{mt}/q_t$ denote the m-th 'output share'. Non-negativity of w_t and y_t means that the elements of $\eta_t = (\eta_{1t}, ..., \eta_{Mt})'$ lie in the unit interval and sum to one. Accordingly, we assume that η_t is independently distributed with Dirichlet pdf[2]

$$p(\eta_t \mid s) = f_D(\eta_t \mid s) \tag{9}$$

where $s = (s_1, ..., s_M)' \in \mathbb{R}_+^M$. Given x_t and the parameters of the weight vector (4), the inverse function theorem says there is a one-to-one mapping between $y_t \in \mathbb{R}_+^M$ and $(\eta_{1t}, ..., \eta_{M-1,t}, q_t)' \in \mathbb{R}_+^M$. The conditional joint density for y_t is therefore

$$p(y_t \mid \alpha, \beta, h, s, u_t) = f_N\left(q_t \mid \delta + b'x_t + x_t'Bx_t - u_t, h^{-1}\right) f_D(\eta_t \mid s) |J_t| \tag{10}$$

---

[1] The notation $f_N(a \mid b, c)$ indicates that a is a normal random variable with mean b and variance c, while the notation $f_G(a \mid b, c)$ indicates that a is a gamma random variable with shape parameter b and scale parameter c (so a has a mean of b/c and a variance of b/c^2).

[2] The notation $f_D(a \mid b)$ is the notation for a Dirichlet pdf used by Poirier (1995, p.132). Other distributional assumptions are possible, including the additive logistic normal.

where $|J_t|$ is the absolute value of the Jacobian of the transformation from $(\eta_{1t}, ..., \eta_{M-1,t}, q_t)$ to y_t; and α is an $M(N + 0.5M + 1.5) \times 1$ vector containing the elements of a, A and H. The Jacobian in this case is

$$J_t = q_t^{1-M} \left(\prod_{m=1}^{M} y_{mt} \right) \left| 0.5A + diag \left(\frac{w_{1t}}{y_{1t}}, ..., \frac{w_{Mt}}{y_{Mt}} \right) \right|. \qquad (11)$$

Finally, the conditional likelihood function for the matrix of observed outputs is

$$p(Y \mid \alpha, \beta, h, s, u) = f_N \left(q \mid X\beta - u, h^{-1}I_T \right) \prod_{t=1}^{T} f_D (\eta_t \mid s) \prod_{t=1}^{T} |J_t|. \qquad (12)$$

This likelihood function has the same structure as one specified by FKS (p.55) and used in a banking application. However, the FKS likelihood function has a different Jacobian term (FKS use a *nonlinear* output aggregator function and the *logarithm* of aggregate output as the dependent variable in the latent stochastic frontier model).

4.2   The Joint Prior

Fernandez et. al. (1997) show that proper priors on the parameters of frontier models are generally needed to ensure the existence of the posterior density. We therefore specify a joint prior of the form $p(\alpha, \beta, h, s, u) = p(\alpha)p(\beta)p(h)p(s)p(u)$ where each of the component priors is proper:

$$p(\alpha) = f_N (\alpha \mid 0_H, k_1 I_L) I (\alpha \in R) \qquad (13)$$

$$p(\beta) = f_N (\beta \mid 0_K, k_2 I_K). \qquad (14)$$

$$p(h) = f_G (h \mid 1, k_3) \qquad (15)$$

$$p(s) = \prod_{m=1}^{M} f_G (s_m \mid 1, k_4) \qquad (16)$$

$$p(u \mid \lambda^{-1}) = \prod_{t=1}^{T} f_G (u_t \mid 1, \lambda^{-1}) \qquad (17)$$

and

$$p(\lambda^{-1}) = f_G (\lambda^{-1} \mid 1, \mu^*) \qquad (18)$$

where K is the row dimension of α, L is the row dimension of β, and R is the region of the parameter space where the constraints $G\nabla_y F (\iota, \iota; g) \leq 0_M$ and $w_t \geq 0_M$ are satisfied. In our empirical example we set $k_1 = k_2 = 10^4$ and $k_3 = k_4 = 10^{-4}$ to ensure the priors for α, β, h and s are relatively noninformative. The pdf (18) is centred on μ^*, a prior estimate of the mean of the inefficiency distribution. In our empirical example we set $\mu^* = 0.2$.

4.3 Posterior Inference

The likelihood function combines with the joint prior to yield a joint posterior density for the unknown parameters and the unobserved inefficiency effects. Analytical integration of this posterior appears impossible, so posterior inference is conducted using MCMC simulation methods. The Gibbs sampling algorithm partitions the vector of unknown parameters and inefficiency effects into blocks, then simulates sequentially from the conditional posterior distribution for each block. Details concerning the algorithm are available in Gelfand and Smith (1990). In the present case, the conditional posteriors are

$$p(\alpha \quad | \quad \beta, h, s, u, \lambda^{-1}, Y) \tag{19}$$

$$\propto \quad \exp\left\{-0.5he'e - 0.5k_1^{-1}\alpha'\alpha\right\} \times \prod_{t=1}^{T} f_D(\eta_t \mid s) \times \prod_{t=1}^{T} |J_t| \times I(\alpha \in R)$$

$$p(\beta \mid \alpha, h, s, u, \lambda^{-1}, Y) = f_N(\beta \mid hVX(q+u), V) \tag{20}$$

$$p(h \mid \alpha, \beta, s, u, \lambda^{-1}, Y) = f_G(h \mid 1 + 0.5T, k_3 + 0.5e'e) \tag{21}$$

$$p(s \mid \alpha, \beta, h, u, \lambda^{-1}, y) = \prod_{t=1}^{T} f_D(\eta_t \mid s) \prod_{m=1}^{M} f_G(s_m \mid 1, k_4) \tag{22}$$

$$p(u \mid \alpha, \beta, h, s, \lambda^{-1}, Y) \propto f_N\left(u \mid X\beta - q - (h\lambda)^{-1}\iota_T, h^{-1}I_T\right) I(u \geq 0_T) \tag{23}$$

and

$$p(\lambda^{-1} \mid \alpha, \beta, h, s, u, Y) = f_G\left(\lambda^{-1} \mid T+1, \mu^* + \iota_T'u\right) \tag{24}$$

where $e \equiv q - X\beta + u$ and $V \equiv \left(hX'X + k_2^{-1}I_K\right)^{-1}$. Simulating from the densities (20), (21), (23) and (24) is straightforward using non-iterative simulation methods. Indeed, simulating from (23) can be accomplished by sampling independently from T univariate truncated normal distributions. Although the remaining densities are nonstandard, they can be simulated using a Metropolis-Hastings algorithm. Details concerning this algorithm can be accessed from Chen et. al. (2000).

5 Empirical Example

As an illustration, we estimate the productive performance of U.S. fossil fuel-fired electricity utilities for a period (1993) just prior to implementation of Phase I of the U.S. acid rain program. Electricity utilities use inputs of labour, capital and fuel to produce outputs of electricity and SO_2. We set $g = (1, -1)'$ to indicate that megawatt hours (MWh) of electricity and tons of SO_2 are desirable and undesirable outputs respectively. The data were assembled by Fare et. al. (2005) from various reports published by the U.S. Energy Information Administration (EIA) and Environmental Protection Agency (EPA). The 209 firms in the sample produced an average of 1.719 million MWh of electricity

and 32,188 tons of the pollutant in 1993. More details concerning the data can be found in Fare et. al.

The Gibbs sampler was used to draw 10^4 observations on the unknown parameters and inefficiency effects. The parameter estimates were constrained so that the output weights were non-negative and monotonicity was satisfied at the variable means, as discussed in Section 4. All programs were written in GAUSS. In this section we report the means, medians and/or standard deviations of (functions of) our MCMC samples. Means and medians of the estimated posterior pdfs are optimal Bayesian point estimates under quadratic loss and absolute loss respectively.

Point estimates of the unknown parameters are presented in Table 1. Since the individual estimates have no meaningful economic interpretations, we simply note that the sample standard deviations associated with the inequality-constrained parameters (elements of α) are generally small relative to their sample means. This is not the case for the unconstrained parameters (elements of β). This illustrates the benefits of incorporating non-sample information into the estimation process when only a few observations are available for estimating a relatively large number of unknown parameters.

Estimates of technical inefficiency effects are of particular interest. The estimated posterior pdf for industry technical inefficiency is depicted in Figure 1, and has a median of 0.216. This indicates that in 1993 it was technically possible for the industry to produce 0.216 additional units of electricity and 0.216 fewer units of SO_2 while holding input levels fixed. Recall that outputs were scaled to have unit means. Thus, our estimate of 0.216 represents $1.719 \times 0.216 = 0.371$ million additional MWh of electricity and $32,188 \times 0.216 = 6,952$ fewer tons of SO_2. By comparison, Fare et al. estimate that firms could have produced an extra 0.320 MWh of electricity and 5,987 fewer tons sulphur dioxide using 1993 input levels.

Efficient firms that operate on the production frontier must forego electricity production in order to reduce emissions of SO_2. The value of the electricity foregone in order to achieve a one ton reduction in SO_2 is given by the shadow price of the pollutant (Fare et. al., 2005):

$$p_{2t} = -p_{1t} \left[\frac{\partial D\left(x_t, y_t; g\right) / \partial y_{2t}}{\partial D\left(x_t, y_t; g\right) / \partial y_{1t}} \right] \tag{25}$$

where p_{1t} and p_{2t} denote the prices of electricity and SO_2 respectively. Only the market price of electricity is observed.

In this paper, we estimate the shadow price of the pollutant by evaluating the right-hand side of equation (25) at the variables means, the only point at which our estimates of the derivatives of the distance function are guaranteed to satisfy the monotonicity property. Our estimate of the shadow price is $6,470 per ton, almost six times higher than the estimate reported by Fare et. al. ($1,117), and almost double that reported by Lee et. al. (2002) ($3,107). Differences in the estimates may be due to the fact that these other authors implicitly assume that the finite sample distributions of their estimators are

Table 1: Parameter Estimates

Parameter	Mean	Median	St. Dev.
a_1	-640.67	-648.49	38.73
a_2	9.45	9.41	1.06
a_{11}	-24.75	-24.79	1.91
a_{12}	23.73	0.21	1.68
a_{22}	4.88	4.83	0.83
h_{11}	-67.61	-68.63	10.25
h_{21}	45.49	46.17	10.30
h_{31}	-33.39	-33.56	6.03
h_{12}	-22.29	-22.37	3.45
h_{22}	-37.24	-37.20	4.77
h_{32}	-6.68	-6.70	2.09
δ	46.41	46.44	47.72
b_1	-7.33	-7.72	38.58
b_2	562.02	564.65	49.84
b_3	10.52	10.48	18.90
b_{11}	-9.19	-9.15	6.69
b_{12}	92.69	93.18	17.91
b_{13}	16.30	16.27	7.08
b_{22}	-23.03	-23.44	13.60
b_{23}	26.17	26.13	8.44
b_{33}	-1.72	-1.72	1.37
h	1.43E-5	1.42E-5	1.73E-6
λ^{-1}	4.55	3.21	4.43
s_1	10.82	10.76	1.34
s_2	1.76	1.75	0.18

symmetric, even though they are only known to be symmetric asymptotically. In contrast, our estimate of the shadow price is the median of an estimated finite-sample distribution which is highly skewed. The estimated posterior pdf for the shadow price of the pollutant is depicted in Figure 2. Observe that the mode of this estimated pdf is approximately $2,500 per ton.

6   Conclusion

Output distance functions representing multi-output technologies can be estimated in a parametric framework. In practice, it is common to algebraically manipulate the distance function into a form that resembles a single-output stochastic production frontier model, and then estimate the model by the method of maximum likelihood. Unfortunately, except in restrictive special cases[3], maximum likelihood estimates are biased due to the fact that outputs that are used as explanatory variables are correlated with the error term. The generalised method of moments can be used to overcome the problem, but finding suitable moment conditions can be difficult. Moreover, the finite sample properties of GMM estimators are unknown.

This paper addresses the endogeneity problem using Bayesian methodology. The approach does not require the identification of moment conditions, and it is possible to draw exact finite sample inferences concerning nonlinear functions of the unknown parameters. As an illustration, the approach was used to estimate the technology of U.S. electricity utilities in 1993. We estimated that inefficient electric utilities could have increased production of electricity and reduced emissions of sulphur dioxide by more than 20% without using more inputs. In addition, we estimated that efficient utilities could have reduced sulphur dioxide emissions at a cost of $6,470 per ton. This estimate is more than double the size of estimates reported elsewhere in the literature (where the endogeneity problem is typically gnored).

One of the aims of this paper was to develop and apply improved methodology for estimating distance functions. The directional output distance function was chosen to explain and illustrate the approach. However, with trivial modification, the approach can also be used to estimate other types of distance functions, including Shephard output and input distance functions.

References

[1] Atkinson, S.W., Cornwell, C. and O. Honerkamp (2003) "Measuring and decomposing productivity change: stochastic distance function estimation versus data envelopment analysis" *Journal of Business and Economic Statistics* 84(April):284-294.

---

[3]For example, when the true (but always unknown) functional form is Cobb-Douglas (Coelli, 2000).
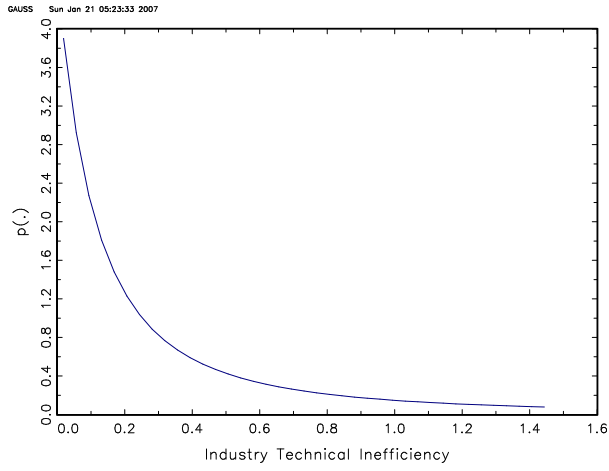
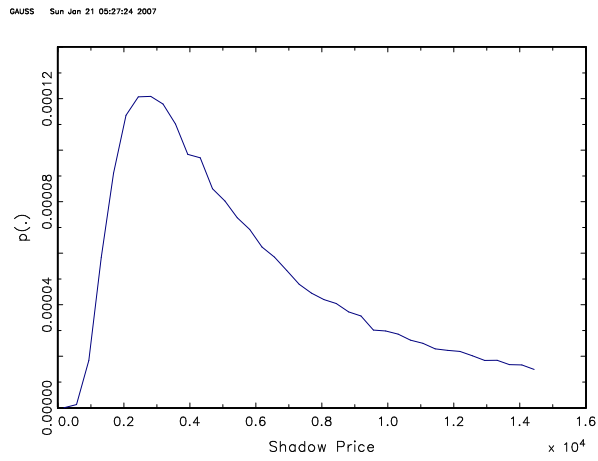Figure 1: Estimated Pdf for Industry Technical Efficiency

Figure 2: Estimated Pdf for Shadow Price of Sulphur Dioxide

11

[2] Atkinson. S.E., R. Färe and D. Primont (1998), "Stochastic Estimation of Firm Inefficiency Using Distance Functions", Working Paper, Department of Economics, University of Georgia, Athens, GA.

[3] Atkinson, S.E., and D. Primont (1998), "Stochastic Estimation of Firm Technology, Inefficiency and Productivity Growth Using Shadow Cost and Distance Functions", Working Paper, Department of Economics, University of Georgia, Athens, GA.

[4] Chen, M.-H., Shao, Q.-M. and J.G. Ibrahim (2000) *Monte Carlo Methods in Bayesian Computation.* Springer. New York.

[5] Coelli, T.J. (2000), "On the Econometric Estimation of the Distance Function Representation of a Production Technology", Discussion Paper 2000/42, Center for Operations Research and Econometrics, Universite Catholique de Louvain.

[6] Coelli, T.J., and S. Perelman (1999), "A Comparison of Parametric and Non-parametric Distance Functions: With Application to European Railways", *European Journal of Operational Research* 117:326-339.

[7] Färe, R., Grosskopf, S., Noh, D.-W. and W. Weber (2005) "Characteristics of a Polluting Technology: Theory and Practice" J*ournal of Econometrics* 126:469-492.

[8] O'Donnell, C.J., and T.J. Coelli (2005), "A Bayesian Approach to Imposing Curvature on Distance Functions", *Journal of Econometrics* 126, 493-523.

[9] Fernandez, C., J. Osiewalski and M.F.J. Steel (1997), "On the Use of Panel Data in Stochastic Frontier Models With Improper Priors" *Journal of Econometrics* 79:169-93.

[10] Fernandez, C., Koop, G. and M. Steel (2005) "Alternative Efficiency Measures for Multiple-Output Production" *Journal of Econometrics* 126, 411-444.

[11] Gelfand, A.E. and A.F.M. Smith (1990) "Sampling-based approaches to calculating marginal densities" *Journal of the American Statistical Association* 85:398-409.

[12] Lee, J-D., Park, J.-B. and T.-Y. Kim (2002) "Estimation of the Shadow Prices of Pollutants With Production/Environment Inefficiency Taken into Account: A Nonparametric Directional Distance Function Approach" *Journal of Environmental Management* 63:365-375.

[13] Poirier, D.J. (1995) *Intermediate Statistics and Econometrics: A Comparative Approach.* MIT Press. London.

[14] Reinhard, S. and G. Thijssen (1998) "Resource Use Efficiency of Dutch Dairy Farms: A Parametric Distance Function Approach" paper presented at Georgia Productivity Workshop III, Athens, Georgia, October 1998.

[15] Robias,D. and C.Arias (2005) "Endogeneity Problems in the Estimation of Multi-output Technologies" mimeo.