Centre for Efficiency and Productivity Analysis

Working Paper Series
No. 03/2003

Title
A Bayesian Approach To Imposing Curvature On Distance Functions

Authors
Chris O'Donnell
Tim Coelli

Date: August 2003

School of Economics
University of Queensland
St. Lucia, Qld. 4072
Australia

A BAYESIAN APPROACH TO IMPOSING CURVATURE ON DISTANCE FUNCTIONS

by

Chris O'Donnell and **Tim Coelli**

School of Economics, University of Queensland

10 August 2003

Abstract

The estimated parameters of output distance functions frequently violate the monotonicity, quasi-convexity and convexity constraints implied by economic theory, leading to estimated elasticities and shadow prices that are incorrectly signed, and ultimately to perverse conclusions concerning the effects of input and output changes on productivity growth and relative efficiency levels. We show how a Bayesian approach can be used to impose these constraints on the parameters of a translog output distance function. Implementing the approach involves the use of a Gibbs sampler with data augmentation. A Metropolis-Hastings algorithm is also used within the Gibbs to simulate observations from truncated pdfs. Our methods are developed for the case where panel data is available and technical inefficiency effects are assumed to be time-invariant. Two models – a fixed effects model and a random effects model – are developed and applied to panel data on 17 European railways. We observe significant changes in estimated elasticities and shadow price ratios when regularity restrictions are imposed.

Keywords: Markov chain Monte Carlo, inequality constraints, output distance function, European railways

1. INTRODUCTION

Multi-input multi-output production technologies that satisfy weak disposability can be described using distance functions. Shephard (1970) and Fare and Primont (1995) discuss both input and output distance functions. An input distance function describes the degree to which a firm can contract its input vector without changing its output vector. An output distance function describes the degree to which a firm can expand its output vector, given an input vector.

The estimation of distance functions has been attracting increasing attention in the efficiency and productivity literature. This interest is most likely due to the fact that distance functions can be used to model multi-input multi-output production technologies without having to aggregate outputs (or inputs), and without having to make behavioural assumptions such as cost-minimisation or profit-maximisation. This is particularly attractive to researchers analysing industries in which public ownership or regulation may make such behavioural assumptions inappropriate.

Distance functions can be estimated using several techniques. Fare, Grosskopf and Lovell (1994) provide a comprehensive discussion of *data envelopment analysis* (DEA), a technique that has the advantage that there is no need to specify a functional form for the boundary of the production technology. Rather, the boundary is constructed using a number of connected hyperplanes, identified by solving a sequence of linear programming problems. However, a downside of DEA is that estimated shadow prices are indeterminate at the intersections of the hyperplanes (though a range of values can be reported), and some may collapse to zero at extreme data points due to the existence of "slack regions".

One way to address this problem is to specify a functional form for the production surface. For example, Fare, Grosskopf, Lovell and Yaisawarng (1993) specify a translog output distance function and estimate it parameters using a generalisation of the linear programming technique proposed by Aigner and Chu (1968). This *parametric linear programming* (PLP) approach identifies a smooth production surface and therefore lends itself to the calculation of shadow prices. However, a drawback of the approach (and also DEA) is that it estimates a deterministic frontier where all deviations from the frontier are implicitly assumed to be due to inefficiency. This means the method is particularly susceptible to the effects of data noise (eg., measurement error), which can lead to biased estimates of the shape and position of the frontier surface.

The issue of data noise was addressed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) for the case of single-output production frontiers. They proposed the estimation of a stochastic production frontier with two error components – a one-sided error term to accommodate inefficiency, and a symmetric error term to accommodate noise. Given the specification of a suitable functional form for the deterministic part of the frontier (eg., translog) and suitable distributional forms for the two error terms (eg., half-normal and normal, respectively), the unknown parameters of the frontier can be estimated using maximum likelihood methods. This *stochastic frontier analysis* (SFA) approach can also be used for distance function estimation – see Coelli and Perelman (1996).

Estimated parametric distance functions have been used for many purposes. For example, they have allowed researchers to measure firm-level technical efficiency by measuring the distance that each firm lies below the production technology – see the analyses of European railways in Coelli and Perelman (1999, 2000). They have also been used to measure and decompose productivity growth through time – see the analysis of Spanish insurance companies by Fuentes, Grifell-Tatjé and Perelman (2001), the analysis of Spanish savings banks by Orea (2002), and the study by Brummer, Glauben and Thijssen (2002) which measures and decomposes productivity growth in dairy farms in Germany, the Netherlands and Poland.

In addition to performance measurement applications, the estimated parameters of distance functions have been used to investigate, for example, the shadow prices of pollutants in electricity generation (Fare, Grosskopf, Lovell and Yaisawarng, 1993; Swinton, 1998), the substitutability of outputs in hospitals (Grosskopf, Margaritis and Valdmanis, 1995), the substitutability of civilian and uniformed personnel in police services (Grosskopf, Hayes and Hirschberg, 1995), and the shadow price of nitrogen pollution in Dutch dairy farms (Reinhard and Thijssen, 1998).

These latter papers extract information on the shadow prices of inputs and/or outputs from the estimated distance functions by exploiting various duality theorems (see Fare and Primont, 1995). This issue is of particular interest to us in this paper because these duality results rely on particular theoretical properties of distance functions. Specifically, they rely on the fact that the output distance function is non-decreasing, convex and homogenous of degree one in outputs, and non-increasing and quasi-convex in inputs[1]; the input distance function is non-increasing, concave and homogenous of degree one in inputs, and non-decreasing and quasi-concave in outputs. We are unaware of any empirical studies in which all these properties have been imposed on parametric (input or output) distance functions. Furthermore, few studies even report the degree to which their estimated functions satisfy these properties. One exception is Reinhard and Thijssen (1998) who report that their estimated output distance function violated convexity in outputs when evaluated at 20.1% of observations, and violated monotonicity (due to an incorrectly signed labour elasticity) when evaluated at 8.6 % of observations.

From a brief survey of distance function applications, all of which involved the use of the translog functional form (or some restricted version of the translog), we observed that all papers had imposed homogeneity, all the PLP papers had imposed monotonicity (ie., the non-increasing/decreasing properties), but no papers had attempted to impose the curvature conditions (ie., the convexity/quasi-convexity and concavity/quasi-concavity properties). This pattern can be explained by the relative ease with which these constraints can be imposed. The homogeneity constraints can be written as linear equality constraints on the parameters and can be easily imposed using either linear programming or econometric methods. The monotonicity constraints are linear inequality constraints which are easy to impose using linear programming, but difficult to impose using traditional econometric approaches, especially since they need to be imposed at each data point. Finally, the curvature constraints are non-linear inequality constraints that also need to be imposed at each data point. This cannot be done using linear programming (the problem must be converted to a non-linear programming problem) and is very difficult using traditional sampling theory econometric methods. While sampling theorists have developed methods for imposing convexity and concavity constraints[2], extension of the methods to deal with quasi-convexity and quasi-concavity is not straightforward – see the discussion in Lau (1978). In theory, Gallant and Gollub's (1984) frequentist method for imposing non-linear constraints can be used, but empirical implementation of the method is complex. In this paper we use a simpler and more intuitively-appealing Bayesian approach.

The relative merits of the Bayesian and sampling theory approaches to inference have been well-articulated by Geweke (1986) and Poirier (1995). For this paper, one of the important advantages of Bayesian methdology is that it allows us to provide exact finite sample results for nonlinear functions of the unknown distance function parameters, including shadow prices and measures of relative technical efficiency. It is also convenient for imposing concavity and convexity constraints, as illustrated by Terrell (1996) and Griffiths, O'Donnell and Tan Cruz (2000) in the context of cost functions. In this paper we extend the Bayesian approach to the imposition of quasi-convexity and quasi-concavity constraints on the parameters of distance functions.

The outline of the paper is as follows. In Section 2 we introduce the output distance function and explain how it can be used to obtain shadow prices. In Section 3 we present the translog output distance function and detail the homogeneity, monotonicity and curvature constraints implied by economic theory. In Section 4 we discuss Bayesian methodology for imposing these constraints on the parameters of the distance function. In Section 5 we apply the methodology to panel data on 17 European railways, and report characteristics of estimated marginal posterior distributions of interest. Section 6 summarises and concludes the paper.

2. OUTPUT DISTANCE FUNCTIONS

We consider the case of a multi-input multi-output production technology where a firm uses the $P \times 1$ input vector $\mathbf{x} = (x_1, ..., x_P)'$ to produce the $M \times 1$ output vector $\mathbf{q} = (q_1, ..., q_M)'$. Following Fare and Primont (1995, p. 8), the production technology can be described by the technology set

$$S = \{(\mathbf{x}, \mathbf{q}) : \mathbf{x} \text{ can produce } \mathbf{q}\}. \tag{2.1}$$

We assume the production technology satisfies a standard set of axioms (Fare and Primont, 1995 p. 27) including convexity, strong disposability,[4] closedness and boundedness.

[1] Fare and Primont (1995, p. 152) state that the output distance function is quasi-concave in inputs and that the input distance function is quasi-convex in outputs. These are typographical errors (personal communication with Dan Primont, 30/9/2002).
[2] These include the Cholesky factorisation approaches discussed in Diewert and Wales (1997) and Ryan and Wales (2000)
[4] Fare and Primont (1995, p. 16) note that an assumption of weak disposability is sufficient for the technology to be fully characterised by the output distance function.

Fare and Primont show that this technology can also be described using an output distance function:[5]

(2.2) $\quad D(\mathbf{x}, \mathbf{q}) = \min\{\delta : \delta > 0, (\mathbf{x}, \mathbf{q}/\delta) \in S\}$.

Our assumptions on the technology set imply that the output distance function is non-decreasing, linearly homogeneous and convex in $\mathbf{q}$, and non-increasing and quasi-convex in $\mathbf{x}$. If $(\mathbf{x}, \mathbf{q})$ belongs to the production set S, then $D(\mathbf{x}, \mathbf{q}) \leq 1$. Moreover, $D(\mathbf{x}, \mathbf{q}) = 1$ if $(\mathbf{x}, \mathbf{q})$ belongs to the "frontier" of the production set. The distance measure $D(\mathbf{x}, \mathbf{q})$ is the inverse of the factor by which the production of all output quantities could be increased while still remaining within the feasible production set, for the given input level. It is equivalent to a Farrell-type output-orientated measure of technical efficiency. See, for example, Fare and Primont (1995, p. 29).

As we have already noted, distance functions are not only used to estimate efficiency levels and productivity change, but are also used to measure shadow prices and the substitution properties of the technology. For example, Grosskopf, Margaritis and Valdmanis (1995) observe that, if the output sets are convex, the duality between the output distance function and the revenue function can be exploited to retrieve information on output shadow prices. Specifically, the partial derivative of the output distance function with respect to the m-th output is a revenue-delated shadow price:

(2.3) $$\frac{\partial D}{\partial q_m} = \frac{p_m^*}{R}$$

where p_m^* is the shadow price of the m-th output and R is total revenue. The ratio of the revenue-deflated shadow prices of two outputs, q_m and q_n,

(2.4) $$\frac{\partial D / \partial q_m}{\partial D / \partial q_n} = \frac{p_m^*}{p_n^*}$$

will reflect the slope of the production possibility curve (ie., the marginal rate of transformation). This ratio can be normalised by the output quantity ratio to obtain a unitless measure of output substitutability.

Similar methods have been employed by Fare *et al* (1993) and Swinton (1998) to calculate shadow prices of pollutants in electricity generation. They use the partial derivatives of the output distance function to construct a ratio as in equation (2.4), and then calculate the shadow price of the pollutant under the assumption that the shadow price of the good output (electricity) equals its observed (market) price.

A similar procedure can be used to extract information on the shadow prices of inputs. Fare and Grosskopf (1994, p. 100) show that the partial derivative of the output distance function with respect to the p-th input provides a measure of the shadow price of the p-th input deflated by total cost (along with a factor reflecting scale economies). Ratios of these partial derivatives (ie. shadow prices) reflect the slope of the isoquant (ie. the marginal rate of technical substitution).

The availability of information on observed prices means that shadow price ratios can be compared with observed price ratios to investigate questions regarding allocative efficiency (in input mix and in output mix). Shadow price information is also utilised in methods that decompose productivity growth into its components. For example, Bremmer *et al* (2000) estimate an output distance function, and then use total differential methods to decompose productivity change into various components, including two allocative-efficiency-related terms that involve differences between shadow shares and observed shares. The shadow share information is obtained by taking derivatives of the estimated (logarithm of the) distance function.

Researchers who derive elasticities and shadow price (or shadow share) information from an estimated distance function need to be confident that the estimated distance function satisfies prescribed monotonicity and curvature properties at each data point in the sample. Monotonicity violations will, for example, give rise to incorrectly-signed elasticities, with the perverse implication that productivity can be improved by increasing inputs while holding outputs fixed. Curvature violations will, for example, give rise to production possibilities frontiers that are convex to the origin, implying the solutions to first-order conditions for revenue maximisation will not be points of maximum revenue. Moreover, shadow prices may be incorrectly signed and may not be unique. These types of results are clearly unsatisfactory. Hence, we seek a method to impose the required regularity conditions on the parameters of an estimated translog distance function.

[5] A more rigorous definition of the output distance function is $D(\mathbf{x}, \mathbf{q}) = \inf\{\delta : \delta > 0, (\mathbf{x}, \mathbf{q}/\delta) \in S\}$.

3. THE TRANSLOG OUTPUT DISTANCE FUNCTION

The translog is a flexible functional form in the sense that it can provide a second-order approximation to an arbitrary functional form. Authors who have used a translog distance function in empirical work include Fare *et al* (1993) (to study electricity generation), Grosskopf, Margaritis and Valdmanis (1995) (hospitals) and Coelli and Perelman (1999) (railways). The translog output distance function defined over M outputs and P inputs can be written as:

$$(3.1) \qquad \ln D = a_0 + \sum_{m=1}^{M} a_m \ln q_m + 0.5 \sum_{m=1}^{M}\sum_{n=1}^{M} a_{mn} \ln q_m \ln q_n + \sum_{p=1}^{P} b_p \ln x_p + 0.5 \sum_{p=1}^{P}\sum_{j=1}^{P} b_{pj} \ln x_p \ln x_j + \sum_{p=1}^{P}\sum_{m=1}^{M} g_{pm} \ln x_p \ln q_m$$

where a_0 and the a_m, a_{mn}, b_p, b_{pj}, and g_{pm} are unknown parameters that satisfy the identifying restrictions $a_{mn} = a_{nm}$ and $b_{pj} = b_{jp}$ for all m, n, j and p. To obtain an empirical version of (3.1) we impose homogeneity constraints and introduce an error term representing statistical noise.

From Euler's Theorem, homogeneity of degree one in outputs implies:

$$(3.2) \qquad \sum_{m=1}^{M} a_m + \sum_{m=1}^{M}\sum_{n=1}^{M} a_{mn} \ln q_n + \sum_{m=1}^{M}\sum_{p=1}^{P} g_{pm} \ln x_p = 1$$

which will be satisfied if[6]

$$(3.3) \qquad \sum_{m=1}^{M} a_m = 1, \qquad \sum_{m=1}^{M} a_{mn} = 0 \ \text{ for all } n, \qquad \text{and} \qquad \sum_{m=1}^{M} g_{pm} = 0 \ \text{ for all } p.$$

Substituting these constraints into the distance function is equivalent to normalising by one of the outputs. If we choose the M-th output, equation (3.1) becomes:

$$(3.4) \qquad \ln(D/q_M) = a_0 + \sum_{m=1}^{M-1} a_m \ln(q_m/q_M) + 0.5 \sum_{m=1}^{M-1}\sum_{n=1}^{M-1} a_{mn} \ln(q_m/q_M)\ln(q_n/q_M) + \sum_{p=1}^{P} b_p \ln x_p$$
$$+ 0.5 \sum_{p=1}^{P}\sum_{j=1}^{P} b_{pj} \ln x_p \ln x_j + \sum_{p=1}^{P}\sum_{m=1}^{M-1} g_{pm} \ln x_p \ln(q_m/q_M)$$

which we write more compactly as:

$$(3.5) \qquad \ln(D/q_M) = TL(\mathbf{x}, \mathbf{q}/q_M, \beta)$$

where $TL(.)$ refers to the translog function and β refers to the vector of a, b and g parameters. An equivalent form of (3.5) is:

$$(3.6) \qquad -\ln q_M = TL(\mathbf{x}, \mathbf{q}/q_M, \beta) + u$$

where $u = -\ln D$ is a non-negative term that captures the effects of inefficiency. If we assume the distance a firm lies from the frontier may be due to either inefficiency or noise, we can follow the stochastic frontier approach proposed by Aigner, Lovell and Schmidt (1977) and Meeusen and van den Broeck (1977) and add a symmetric error term, v, to capture the effects of data noise. The resulting empirical output distance function is

$$(3.7) \qquad -\ln q_M = TL(\mathbf{x}, \mathbf{q}/q_M, \beta) + v + u.$$

which is in the form of a standard stochastic frontier model. In Section 4 we describe how to use panel data, and the assumption that the u terms are invariant across time, to estimate the parameters of this model under both fixed and random effects assumptions. Our objective will be to estimate the parameters of the models in such a way that the estimated functions satisfy the following monotonicity (ie. non-increasing in **x** and non-decreasing in **q**) and curvature (ie. quasi-convex in **x** and convex in **q**) properties implied by production theory.

---

[6] Although these conditions are both necessary and sufficient for global homogneity of the distance function (3.1), they are sufficient but *not* necessary for global homogeneity of the true (unknown) distance function. This implies the parameters of our flexible functional form may be overconstrained. This issue is addressed in a broader context by O'Donnell (1999).

Monotonicity and Curvature Constraints

Monotonicity and curvature conditions involve constraints on functions of the partial derivatives of the distance function. Key derivatives are the elasticities of distance with respect to inputs and outputs:

$$(3.8) \qquad s_p \equiv \frac{\partial \ln D}{\partial \ln x_p} = b_p + \sum_{j=1}^{P} b_{pj} \ln x_j + \sum_{m=1}^{M} g_{pm} \ln q_m$$

and

$$(3.9) \qquad r_m \equiv \frac{\partial \ln D}{\partial \ln q_m} = a_m + \sum_{n=1}^{M} a_{mn} \ln q_n + \sum_{p=1}^{P} g_{pm} \ln x_p.$$

For D to be non-increasing in **x** we require:

$$(3.10) \qquad f_p \equiv \frac{\partial D}{\partial x_p} = \frac{\partial \ln D}{\partial \ln x_p} \frac{D}{x_p} = s_p D / x_p \le 0 \;\Leftrightarrow\; s_p \le 0$$

while for D to be non-decreasing in **q** we require:

$$(3.11) \qquad h_m \equiv \frac{\partial D}{\partial q_m} = \frac{\partial \ln D}{\partial \ln q_m} \frac{D}{q_m} = r_m D / q_m \ge 0 \;\Leftrightarrow\; r_m \ge 0.$$

For quasi-convexity in **x**, we arrange the first- and second-order derivatives of D to form the bordered Hessian matrix:

$$(3.12) \qquad \mathbf{F} = \begin{bmatrix} 0 & f_1 & \ldots & f_P \\ f_1 & f_{11} & \ldots & f_{1P} \\ : & : & \ldots & : \\ f_P & f_{1P} & \ldots & f_{PP} \end{bmatrix}$$

where

$$(3.13) \qquad f_{pj} \equiv \frac{\partial^2 D}{\partial x_p \partial x_j} = \frac{\partial f_p}{\partial x_j} = \frac{\partial (s_p D / x_p)}{\partial x_j} = (b_{pj} + s_p s_j - \delta_{pj} s_p)(D / x_p x_j)$$

and $\delta_{pj} = 1$ if $p = j$ and 0 otherwise. For D to be quasi-convex in **x** over the nonnegative orthant it is sufficient[7] that all the principal minors of **F** be negative (Chiang, 1984, p.394).

Finally, for convexity in **q** we form the Hessian matrix

$$(3.14) \qquad \mathbf{H} = \begin{bmatrix} h_{11} & h_{12} & \ldots & h_{1M} \\ h_{12} & h_{22} & \ldots & h_{2M} \\ : & : & \ldots & : \\ h_{1M} & h_{2M} & \ldots & h_{MM} \end{bmatrix}$$

where

$$(3.15) \qquad h_{mn} \equiv \frac{\partial^2 D}{\partial q_m \partial q_n} = \frac{\partial h_m}{\partial q_n} = \frac{\partial (r_m D / q_m)}{\partial q_n} = (a_{mn} + r_m r_n - \delta_{mn} r_m)(D / q_m q_n).$$

[7] This sufficient condition will be used to impose quasi-convexity in our empirical work. To check for quasi-convexity violations in unconstrained models, we check a necessary condition for quasi-convexity, namely that all the principal minors of **F** be non-positive.

The function D will be convex in $\mathbf{q}$ over the nonnegative orthant if and only if $\mathbf{H}$ is positive semidefinite (Lau, p.414). Thus, D will be convex in $\mathbf{q}$ if and only if all the principal minors of $\mathbf{H}$ are non-negative (Rao and Bhimasankaram, 1992, p.344). In the application to European railways discussed below we have $M = 2$ outputs. In this special case, the Hessian matrix will be positive semidefinite if and only if $a_{11} \geq r_1 r_2$ (≤ 0.25)[8].

4. BAYESIAN ESTIMATION WITH PANEL DATA

This section draws on the discussions of Bayesian stochastic frontier models found in the papers by Koop, Osiewalski and Steel (1997) and Koop and Steel (2001). We introduce some slightly different notation to avoid confusion with the notation we have already used to describe distance functions in Sections 2 and 3, and to avoid having to redefine variables when we switch the discussion from the fixed effects model to the random effects model.

We assume data is available for $i = 1, .., N$ firms for $t = 1, ..., T$ time periods and write the stochastic output distance function model from equation (3.7) as:

(4.1) $\quad y_{it} = a_0 + \mathbf{z}_{it}'\boldsymbol{\phi} + v_{it} + u_i$

where $y_{it} = -\ln q_{Mit}$, $\mathbf{z}_{it}$ is a $K \times 1$ vector comprising functions of the logarithms of the input variables and the output ratios, $\boldsymbol{\phi}$ is a $K \times 1$ vector of parameters, and a_0 is the intercept parameter in (3.1) and (3.4). The set of T observations on firm i can be written:

(4.2) $\quad \mathbf{y}_i = a_0\mathbf{j}_T + \mathbf{Z}_i\boldsymbol{\phi} + \mathbf{v}_i + u_i\mathbf{j}_T$

where $\mathbf{j}_T$ is a $T \times 1$ unit vector, both $\mathbf{y}_i = (y_{i1}, ..., y_{iT})'$ and $\mathbf{v}_i = (v_{i1}, ..., v_{iT})'$ are $T \times 1$, and $\mathbf{Z}_i$ is $T \times K$. For estimation purposes we assume the elements of the $\mathbf{v}_i$s are independent normal random variables with zero means and constant variance, h^{-1}. The probability density function (pdf) of $\mathbf{v}_i$ is written[9] $p(\mathbf{v}_i \mid h) = f_N(\mathbf{v}_i \mid \mathbf{0}_T, h^{-1}\mathbf{I}_T)$.

We can estimate the model assuming the time-invariant u_i terms are either fixed parameters or random variables. In either case, the technical efficiency of firm i is

(4.3) $\quad D_i = TE_i = \exp(-u_i)$

For small u_i, $u_i \approx 1 - \exp(-u_i) = 1 - D_i$, so u_i can sometimes be used as a measure of technical *in*efficiency (Kim and Schmidt, 2000, p.93).

The Fixed Effects Model

If the u_i terms in (4.1) and (4.2) are treated as fixed parameters then the so-called *fixed effects model* can be written:

(4.4) $\quad \mathbf{y}_i = \alpha_i\mathbf{j}_T + \mathbf{Z}_i\boldsymbol{\phi} + \mathbf{v}_i$

where $\alpha_i = a_0 + u_i$ is the i-th individual effect. This is the form of a standard panel data model (eg. Judge *et al*, 1985, p.519). The vector of individual effects is $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_N)'$, and the complete set of NT observations can be written compactly in the form

(4.5) $\quad \mathbf{y} = (\mathbf{I}_N \otimes \mathbf{j}_T)\boldsymbol{\alpha} + \mathbf{Z}\boldsymbol{\phi} + \mathbf{v} = \mathbf{W}\boldsymbol{\theta} + \mathbf{v}$

where $\mathbf{y} = (\mathbf{y}_1', ..., \mathbf{y}_N')'$ and $\mathbf{v} = (\mathbf{v}_1', ..., \mathbf{v}_N')'$ are $NT \times 1$, $\mathbf{Z} = (\mathbf{Z}_1', ..., \mathbf{Z}_N')'$ is $NT \times K$, $\mathbf{W} = (\mathbf{I}_N \otimes \mathbf{j}_T, \mathbf{Z})$ is $NT \times (N + K)$ and $\boldsymbol{\theta} = (\boldsymbol{\alpha}', \boldsymbol{\phi}')'$ is $(N + K) \times 1$.

---

[8] Homogeneity requires $a_{11} = -a_{12} = -a_{21} = a_{22}$ and $r_1 + r_2 = 1$, and these constraints can be used to show that $|\mathbf{H}| = 0$. Then $\mathbf{H}$ will be psd if and only if $(a_{11} + r_1 r_1 - r_1)(D/q_1 q_1) \geq 0$, or $a_{11} \geq r_1 r_2$. Monotonicity and homogeneity together imply $r_1 r_2 \leq 0.25$.

[9] The notation $f_N(\mathbf{a} \mid \mathbf{b}, \mathbf{C})$ is used to indicate that $\mathbf{a}$ is a multivariate normal random vector with mean vector $\mathbf{b}$ and covariance matrix $\mathbf{C}$. The notation $f_G(a \mid b, c)$ will be used to indicate that a has a Gamma distribution with shape parameter b and scale parameter c (so a has mean b/c and variance b/c^2).

For Bayesian inference we adopt the following independent priors for the unknown h and θ:

(4.6) $p(h) \propto h^{-1}$

and

(4.7) $p(\theta) \propto I(\theta \in R_j)$

where $I(.)$ is an indicator function which takes the value one if the argument is true and zero otherwise, and R_j is the set of permissible parameter values when no constraints (j = 0), monotonicity constraints (j = 1) and both monotonicity and curvature constraints (j = 2) must be satisfied. The joint prior pdf is therefore

(4.8) $p(\theta, h) \propto h^{-1} \times I(\theta \in R_j)$.

The likelihood function is

(4.9) $p(\mathbf{y} \mid \theta, h) = f_N(\mathbf{y} \mid \mathbf{W}\theta, h^{-1}\mathbf{I}_{NT}) \propto h^{NT/2} \times \exp\{-0.5h\, [\mathbf{y} - \mathbf{W}\theta]'[\mathbf{y} - \mathbf{W}\theta]\}$

and, using Bayes's Theorem, the posterior pdf is

(4.10) $p(\theta, h \mid \mathbf{y}) \propto p(\mathbf{y} \mid \theta, h)\, p(\theta, h) \propto h^{NT/2\ -\ 1} \times \exp\{-0.5h\, [\mathbf{y} - \mathbf{W}\theta]'[\mathbf{y} - \mathbf{W}\theta]\} \times I(\theta \in R_j)$.

We are primarily interested in θ so we integrate h out of this pdf to obtain

(4.11) $p(\theta \mid \mathbf{y}) \propto \{(\mathbf{y} - \mathbf{W}\theta)'(\mathbf{y} - \mathbf{W}\theta)\}^{-NT/2} \times I(\theta \in R_j)$.

When j = 0 the prior (4.7) is an unconstrained uniform prior and the posterior (4.11) is in the form of a multivariate-t distribution. In this case, characteristics of the marginal posterior densities of elements of θ can be obtained using standard Bayesian results (eg. Zellner, 1971, pp.66-70). However, when $j > 0$ the vector θ will have a truncated joint posterior and MCMC simulation methods are required to estimate characteristics of the marginal posteriors. Indeed, irrespective of the value of j (ie. whether or not θ is constrained), MCMC methods are required to estimate characteristics of the marginal posteriors of shadow price ratios and other economic quantities of interest, denoted $g(\theta)$. Accordingly, for both unconstrained and constrained models, we use MCMC methods to draw sample observations $\{\theta^j: j = 1, ..., J\}$ from the posterior $p(\theta \mid \mathbf{y})$. Integrals of the form

(4.12) $\mathrm{E}\{g(\theta) \mid \mathbf{y}\} = \int g(\theta)\, p(\theta \mid \mathbf{y}) d\theta$

are then estimated by simply averaging $g(\theta)$ over these J draws.

Simulating from the unconstrained (ie., j = 0) version of the posterior (4.11) is possible using a basic Gibbs sampler. The Gibbs sampling algorithm draws from the joint posterior density by sampling from a series of conditional posteriors. Details of the algorithm are available in the seminal paper by Gelfand and Smith (1990). However, to impose monotonicity and curvature restrictions the basic Gibbs sampler needs to be supplemented with an accept-reject algorithm. Terrell (1996) used this approach to impose monotonicity and concavity constraints on the parameters of a cost function. A disadvantage of the Terrell approach is that it may be necessary to generate an extremely large number of candidate draws before finding one that is permissible. In many situations, a more efficient alternative is to simulate from the constrained posterior using a Metropolis-Hastings (M-H) algorithm. Details of the M-H algorithm can be found in, for example, Chen, Shao and Ibrahim (2000, p.23-24). Random-walk M-H algorithms have been used by O'Donnell, Shumway and Ball (1999) and Griffiths, O'Donnell and Tan Cruz (2000) to impose curvature constraints on the parameters of systems of cost and/or conditional input demand functions.

Implementation of the random-walk M-H algorithm involves chooosing an arbitrary proposal density to generate candidates for inclusion in the MCMC sequence. In our empirical application we use a multivariate normal proposal density, with covariance matrix equal to a tuning scalar multiplied by the maximum likelihood estimate of the covariance matrix of the parameters. The tuning scalar is used to manipulate the acceptance rate (ie., the rate at which candidate draws are included in the MCMC sample). Roberts, Gelman and Gilks (1997) show that if the target and proposal densities are normal pdfs, the optimal acceptance rate (ie., the one which minimises the autocorrelations across the sample values) is between 0.45 (in one-dimensional problems) and approximately 0.23 (as the number of dimensions becomes infinitely large). In our empirical work we choose the tuning scalar so that the acceptance rate lies within this range.

Our simulated draws from the constrained and unconstrained posteriors are used to estimate characteristics of the marginal pdfs of functions of the parameters, including measures of relative technical efficiency. The fixed effects model cannot generally be used to estimate the technical efficiency scores given by (4.3). However, it can be used to estimate *relative* technical efficiencies given by

$$(4.13) \quad RTE_i = \exp(\min_j(\alpha_j) - \alpha_i) = \exp(\min_j(u_j) - u_i).$$

These RTE_i will lie between zero and one, with a value of one indicating the firm is the "best" firm in the sample. If both N and T are large, they can be regarded as absolute technical efficiency scores (Kim and Schmidt, 2000, p.96).

The Random Effects Model

Rather than assume the u_i's in (4.1) and (4.2) are fixed, we now assume they are independent random variables. The so-called *random effects model* can be written:

$$(4.14) \quad \mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{v}_i + u_i\mathbf{j}_T$$

where $\mathbf{X}_i = (\mathbf{j}_T, \mathbf{Z}_i)$ is $T \times (K + 1)$ and $\boldsymbol{\beta} = (a_0, \boldsymbol{\phi}')'$ is $(K + 1) \times 1$. The complete set of NT observations can be compactly written:

$$(4.15) \quad \mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{v} + (\mathbf{I}_N \otimes \mathbf{j}_T)\mathbf{u}$$

where $\mathbf{X} = (\mathbf{X}_1', ..., \mathbf{X}_N')'$ is $NT \times (K + 1)$ and $\mathbf{u} = (u_1, ..., u_N)'$ is $N \times 1$. In this paper we assume the elements of $\mathbf{u}$ are independent random variables drawn from exponential distributions that share a common unknown parameter, λ. Specifically, we assume $p(u_i \mid \lambda^{-1}) = f_G(u_i \mid 1, \lambda^{-1})$. More flexible distributions (eg. gamma with fixed shape parameter greater than one) can be easily handled in the Bayesian framework (for examples, see van den Broeck *et al*, 1994; Koop *et al*, 1995) but they make the inefficiency errors more difficult to distinguish from the normally distributed errors representing noise (see Ritter and Simar, 1997). Another reason for choosing the exponential distribution is that van den Broeck *et al* (1994) find that models based on this distribution are reasonably robust to changes in priors.

We adopt the independent prior (4.6) for h and the following prior for $\boldsymbol{\beta}$:

$$(4.16) \quad p(\boldsymbol{\beta}) \propto I(\boldsymbol{\beta} \in R_j)$$

Fernandez, Osiewalski and Steel (1997) show that we need a proper prior for the remaining parameter, λ, in order to obtain a proper posterior. Accordingly, we use the proper prior

$$(4.17) \quad p(\lambda^{-1}) = f_G(\lambda^{-1} \mid 1, -\ln(\tau^*))$$

where τ^* is the prior median of the efficiency distribution. Our joint prior pdf is therefore

$$(4.18) \quad p(\boldsymbol{\beta}, h, \mathbf{u}, \lambda^{-1}) = p(\boldsymbol{\beta})\, p(h)\, p(\mathbf{u} \mid \lambda^{-1})\, p(\lambda^{-1}) \propto h^{-1} \times I(\boldsymbol{\beta} \in R_j) f_G(\lambda^{-1} \mid 1, -\ln(\tau^*)) \times \prod_{i=1}^{N} f_G(u_i \mid 1, \lambda^{-1})$$

Koop, Steel and Osiewalski (1995) set $\tau^* = 0.875$ in their study of efficiency in the US electric utility industry. By coincidence, our best prior knowledge of the efficiency of European railways in this study is the mean efficiency value of 0.878 reported by Coelli and Perelman (2000). Koop *et al* (1997) find their results for US hospitals are extremely robust to enormous changes in τ^*, so we are comfortable setting $\tau^* = 0.878$.

The likelihood function is

$$(4.19) \quad p(\mathbf{y} \mid \boldsymbol{\beta}, h, \mathbf{u}, \lambda^{-1}) \propto h^{NT/2} \exp\{-(h/2)[\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - (\mathbf{I}_N \otimes \mathbf{j}_T)\mathbf{u}]'[\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - (\mathbf{I}_N \otimes \mathbf{j}_T)\mathbf{u}]\}$$

and the posterior $p(\beta, h, \mathbf{u}, \lambda^{-1} | \mathbf{y})$ is proportional to the product of (4.18) and (4.19).

To draw observations from the posterior it is convenient to use a Gibbs sampler with data augmentation. The term "data augmentation" derives from the fact that it is convenient to augment the observed data by drawing observations on **u**. Such an algorithm has been used to estimate unconstrained stochastic frontier models by several authors including Koop, Steel and Osiewalski (1995).

The Gibbs sampler with data augmentation involves drawing sequentially from the following conditional posteriors:

$$(4.20) \quad p(\lambda^{-1} \mid \mathbf{y}, \beta, h, \mathbf{u}) \propto f_G(\lambda^{-1} \mid N+1, \mathbf{u}'\mathbf{j}_N - \ln(\tau^*))$$

$$(4.21) \quad p(h \mid \mathbf{y}, \beta, \mathbf{u}, \lambda^{-1}) \propto f_G(h \mid NT/2, 0.5[\mathbf{y} - \mathbf{X}\beta - (\mathbf{I}_N \otimes \mathbf{j}_T)\mathbf{u}]'[\mathbf{y} - \mathbf{X}\beta - (\mathbf{I}_N \otimes \mathbf{j}_T)\mathbf{u}])$$

$$(4.22) \quad p(\beta \mid \mathbf{y}, h, \mathbf{u}, \lambda^{-1}) \propto f_N(\beta \mid \mathbf{b}, h^{-1}(\mathbf{X}'\mathbf{X})^{-1}) \times I(\beta \in R_j)$$

and

$$(4.23) \quad p(\mathbf{u} \mid \mathbf{y}, \beta, h, \lambda^{-1}) \propto f_N(\mathbf{u} \mid \bar{\mathbf{y}} - \bar{\mathbf{X}}\beta - (Th\lambda)^{-1}\mathbf{j}_N, (Th)^{-1}\mathbf{I}_N) \times \prod_{i=1}^{N} I(u_i \geq 0)$$

where $\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'[\mathbf{y} - (\mathbf{I}_N \otimes \mathbf{j}_T)\mathbf{u}]$ is $(K + 1) \times 1$,

$\bar{\mathbf{X}} = (\bar{\mathbf{x}}_1', ..., \bar{\mathbf{x}}_N')'$ is $N \times (K + 1)$,

$\bar{\mathbf{x}}_i = (1/T)\mathbf{j}_T'\mathbf{X}_i$ is $1 \times (K + 1)$,

$\bar{\mathbf{y}} = (\bar{\mathbf{y}}_1, ..., \bar{\mathbf{y}}_N)'$ is $N \times 1$,

and $\bar{\mathbf{y}}_i = (1/T)\mathbf{j}_T'\mathbf{y}_i$ is a scalar.

Draws from these conditional posteriors will converge to draws from the posterior $p(\beta, h, \mathbf{u}, \lambda^{-1} | \mathbf{y})$. In practice, sampling from the gamma distributions (4.20) and (4.21) is straightforward. Sampling from the truncated multivariate normal distribution (4.22) can be accomplished using a simple accept-reject algorithm, but in our empirical application we found it was more efficient to use the M-H algorithm discussed earlier in the context of the fixed effects model. In the case of a (non-truncated) tri-variate normal distribution, Gelfand and Lee (1993) find that an appropriate number of M-H sub-iterations at each stage of an outer MCMC algorithm is 4 or 8. In this paper we use 10 M-H sub-iterations at each stage of the outer Gibbs algorithm..

Finally, sampling from the conditional density (4.23) is simplified by noting that the covariance matrix is a scalar times an identity matrix, and the truncations are independent. Thus, we sample from this multi-dimensional conditional posterior using N univariate truncated normal distributions, using results that can be found in Albert and Chib (1996).

Once again, we are interested in characteristics of the marginal pdfs of (functions of) θ, including the measure of absolute technical efficiency given by (4.3). Again, MCMC draws from the posterior $p(\beta, h, \mathbf{u}, \lambda^{-1} | \mathbf{y})$ can be used to compute corresponding draws on these quantities of interest.

5. APPLICATION TO EUROPEAN RAILWAYS

Data

Our data are observations on 17 European railways over the six-year period from 1988 to 1993. The data are derived from data published by the International Union of Railways (UIC, 1988-1993). Our model is defined with two output variables (passengers and freight) and three input variables (labour, rolling stock and lines)[10].

[10] Information on energy use was not available. It is expected that energy use would be closely correlated with rolling stock and hence its omission is unlikely to introduce serious bias.

The passenger service output and freight service output variables are measured using the sum of distances travelled by each passenger and the sum of distances travelled by each tonne of freight, respectively.

The labour input variable is measured by the annual mean of monthly data on staff levels. These staff levels only relate to those staff involved in train services and station services. Staff involved in the maintenance of rolling stock and lines are not included given that some companies subcontract these activities.[11] Rolling stock is measured by the sum of available freight wagons and coach transport capacities in tonnes and seats, respectively. The third input used is measured using the total length of lines.[12] More details concerning the construction of the data are provided in the Appendix.

Empirical results

The first step in our empirical analysis was to estimate the fixed-effects model using least squares methods. Following estimation, we checked the monotonicity and curvature conditions described in Section 3 at each data point in the sample of 102 observations. We found the monotonicity constraints were violated at 67 data points (8, 37 and 29 violations in the cases of inputs 1, 2 and 3; no violations for output 1; 5 violations for output 2) and the curvature constraints were violated at every data point (quasi-convexity in inputs was not satisfied at any data point; convexity in outputs was satisfied everywhere). Thus, the unrestricted estimated output distance function fails to satisfy the properties prescribed by production theory, and any measures derived from this estimated function, such as elasticities or shadow prices, are likely to be unreliable.

The next step was to estimate the model using the fixed effects and random effects Bayesian methods described in the previous section. These models were estimated under three different levels of constraints: no constraints ($j = 0$); monotonicity only ($j = 1$); and monotonicity and curvature ($j = 2$). In each case we generated a total of 60,000 observations, and then discarded the first 10,000 as a 'burn-in'. The means of our six MCMC samples are estimates of the means of the marginal posterior distributions of the parameters and are reported in Tables 1 and 2. These tables also report 95% posterior coverage regions calculated as the fifth and ninety-fifth percentiles of the MCMC sample observations[13]. There is broad similarity between the estimates obtained from the fixed and random effects models – the fixed effects point estimates are contained within the random effects coverage regions and vice versa; if anything, the fixed effects coverage regions tend to be slightly wider than those obtained using the random effects model.

Prior to estimation the sample data was deflated so that each variable had a sample mean of one. The derivatives (3.8) and (3.9) collapse to b_p and a_m when evaluated at these (unit) variable means, and the monotonicity conditions can therefore be expressed as $b_p \leq 0$ and $a_m \geq 0$. The point estimate of b_2 reported in the first column of Table 2 (the unconstrained fixed effects model) is positive, implying the theoretical monotonicity constraints are not satisfied. The estimates of the first-order coefficients reported in the fourth column of Table 2 (the unconstrained random effects model) are correctly signed.[14]

Further insights into regularity violations can be gained by examining the coverage regions and (estimated) marginal posterior pdfs. For example, even though the point estimate of b_2 reported in the fourth column of Table 2 is correctly signed, the associated coverage region spans zero, meaning there is positive probability that monotonicity is violated. The associated unconstrained marginal posterior pdf is depicted in Figure 1 and reveals that b_2 is incorrectly signed with estimated probability 0.42 (the estimated area under the pdf and to the right of zero). Further analysis of the estimated unconstrained marginal posterior pdfs of the s_p and r_m and the principal minors of $\mathbf{F}$ and $\mathbf{H}$ at points other than the variable means reveals that there is positive probability that monotonicity is violated at every data point in the sample[15] (47, 67 and 56 data points in the cases of inputs 1, 2 and 3; 16 and 15 in the cases of outputs 1 and 2). Moreover, there is positive probability that curvature is also violated at every data point, all due to failure of the unconstrained estimated distance function to satisfy the property of quasi-convexity in inputs.

---

[11] For instance, the Swedish railways infrastructure is an independent company (BV).

[12] For a discussion of alternative railways technology specifications, see Cowie and Riddington (1996).

[13] These coverage regions are analogous to the 95% confidence intervals used by frequentists. We report coverage regions instead of estimated standard deviations because they provide a better indication of likely and unlikely values of the parameters in cases like ours where many of the marginal posterior distributions are asymmetric.

[14] Our use of a noninformative joint prior means the unconstrained Bayesian fixed-effects results are quite similar to those obtained using least squares. Hence we have not reported the least squares results here.

[15] Using results from the fixed effects model.

It is apparent from Table 2 and Figure 1 that incorporating monotonicity information into the estimation process has had the effect of dramatically reducing the variances of the estimated marginal pdfs. This is intuitively plausible, and consistent with the Monte Carlo finding of Dorfman and McIntosh (2001) that imposing inequality restrictions on systems of demand equations can improve the mean squared errors (MSEs) on estimated elasticities by up to 50 percent. However, monotonicity constraints alone are not enough to ensure the estimated distance function is regular. Monotonicity-constrained point estimates of the s_p and r_m and the principal minors of **F** and **H** reveal that curvature is still violated at every data point in the sample, although quasi-convexity in inputs is now violated at only 30 data points (instead of 102 using the unconstrained estimates), and convexity in outputs is violated at 100 data points (instead of none). When we focus on the estimated posterior pdfs, we find there is positive probability that quasi-convexity in inputs is violated at 98 data points (instead of 102 using the unconstrained estimates), and convexity in outputs is violated at 100 data points (instead of none)[16]. Imposing quasi-convexity and convexity has had relatively little impact on the signs and magnitudes of our estimates of the first-order coefficients and their standard errors. However, imposing these properties has had a relatively large impact on some of the second-order coefficients – point estimates of b_{13} and b_{23} undergo sign reversals, and we observe a ten-fold decrease in the width of some coverage regions (eg., that for of a_{11}).

Point estimates of (time-invariant) relative technical efficiencies are reported in Table 3. The unconstrained estimates obtained using the fixed effects model are generally lower than the random effects estimates, reflecting our use of the uniform prior (4.7). This prior implies a noninformative prior for $\min_j(u_j) - u_i$ on the interval $[0, \infty)$, and in turn this implies an informative prior on $RTE_i = \exp(\min_j(u_j) - u_i)$ of the form $p(RTE_i) \propto 1/RTE_i$ (Kim and Schmidt, p. 103). This is an L-shaped prior which is strongly biased in favour of low efficiency, and since we have only $T = 6$ observations per firm, it is a prior that is unlikely to be dominated by the data[17]. Kim and Schmidt (2000) observe a similar phenomenon in the case of Indonesian rice farms and Texas utilities. A second noteworthy feature of Table 3 is that the random effects estimates tend to exhibit less variation across firms than the fixed effects estimates. This reflects our assumption that the random inefficiency effects are drawn from a common distribution, ie. $f_G(u_i \mid 1, \lambda^{-1})$. The fixed effects estimates of technical efficiencies tend to fall as we impose monotonicity and curvature constraints, whereas the random effects estimates are relatively robust to the imposition of these regularity conditions.

Estimated distance functions are not just used to estimate efficiency effects – they are also used to obtain shadow price information and to calculate and decompose productivity growth measures. Such work utilises estimates of the partial derivatives of the estimated distance function. In order to illustrate the effects of imposing monotonicity and curvature constraints on (functions of) these partial derivatives, we have reported a selection of estimates in Tables 4 to 7 and Figures 2 to 9. In Tables 4 to 6 we report information on input elasticities (s_p) evaluated at the input and output levels of each firm in the final year of the sample period (1993). In Table 7 we report corresponding information on the output shadow price ratios (passengers over freight) given by (2.4). Figures 2 to 9 present corresponding estimated marginal pdfs for firm $i = 3$ only (Swiss Federal Railways).

It is apparent from Tables 4 to 6 that several (estimated) unconstrained elasticities are incorrectly signed (ie. are positive) – two incorrectly-signed estimates in the case of input 1, eight in the case of input 2, and five in the case of input 3. The larger number of violations in the case of the elasticity for input 2 is largely due to the fact that the estimated unconstrained marginal pdf of b_2 extends a long way into the positive domain (see Table 2 and Figure 1). Incorrectly-signed elasticity estimates imply (implausible) negative shadow price estimates. They also lead to perverse conclusions concerning productivity growth – a positive elasticity implies that an increase in the use of that input (with all other variables, including output, held constant) will increase the (measured) productivity of that firm.

Further inspection of Tables 4 to 6 reveals that all the monotonicity-constrained estimates are correctly signed (by construction). The effects on the elasticity estimates of subsequently imposing curvature constraints are also quite noticeable. This is clearly illustrated in Figures 2 to 7 in the case of firm 3 – the curvature constraints cause rightward and leftward shifts in the estimated pdfs for the elasticities of inputs 1 and 3; the variances of the estimated pdfs become smaller as more inequality information is incorporated into the estimation process, in line with our findings concerning the estimated pdfs of individual coefficients (see Figure 1).

Our final comments concern the estimated (distributions of the) output shadow price ratios presented in Table 7 and Figures 8 and 9. Monotonicity violations are evident among the unconstrained estimates – the first and

[16] Again, using results from the fixed effects model.
[17] By implication, our results may not be robust to our choice of prior.

fourth columns of Table 7 contain negative point estimates of shadow price ratios for a couple of firms, and a small number of coverage regions span zero[18]; Figures 8 and 9 depict estimated unconstrained marginal pdfs that extend into the negative domain. The monotonicity-constrained estimates are correctly signed by construction and, once again, the estimated variances of the marginal posterior pdfs decrease with the imposition of (more) regularity constraints. The estimated shadow price ratios reported in Table 7 can be used for several purposes. For example, they could be compared with observed price ratios to investigate the degree of cross-subsidisation between freight and passenger services in this industry. It is clear that findings from these types of investigations will be sensitive to the imposition of monotonicity and curvature constraints.

5. SUMMARY AND CONCLUSION

The estimation of output distance functions is popular among applied microeconomists, partly because distance functions obviate the need to make behavioural assumptions (eg., cost minimisation) in order to estimate interesting characteristics of multi-input multi-output technologies. Characteristics of particular interest to economists include elasticities which measure the effects on efficiency of changes in inputs, and shadow price ratios which measure marginal rates of transformation between outputs, or marginal rates of technical substitution between inputs. To recover reliable estimates of these characteristics, it is important to estimate distance functions in a manner consistent with the regularity (eg., homogeneity , quasi-convexity and convexity) properties implied by economic theory. Failure to impose these regularity constraints on the parameters of distance functions may give rise to estimated elasticities and shadow price ratios that are unreliable, if not implausible.

Sampling theorists seem to have little difficulty imposing monotonicity and convexity constraints[19] on the parameters of distance functions, but imposing quasi-convexity constraints appears difficult. This paper handles the problem in a Bayesian framework. Imposition of monotonicity and curvature constraints is straightforward in a Bayesian framework, but involves the use of Markov chain Monte Carlo (MCMC) simulation techniques. This paper uses two common MCMC algorithms to estimate the parameters of a translog output distance function under two different assumptions concerning inefficiency effects. We use a Metropolis-Hastings algorithm to estimate the distance function under a fixed effects assumption. We use a Gibbs sampler with data augmentation and Metropolis-Hastings sub-chains to estimate the distance function under a random effects assumption. By imposing monotonicity, quasi-convexity and convexity constraints on the parameters of an output distance function, we extend earlier work by Koop, Steel and Osiewalski (1995) and Koop, Osiewalski and Steel (1997) on Bayesian estimation of unconstrained stochastic production frontiers.

In our empirical application to 17 European railways, our estimates of (relative) technical efficiency seem more sensitive to the random versus fixed effects assumptions than to the imposition of regularity constraints. However, the imposition of these constraints gives rise to significant changes in the signs and magnitudes of other estimated functions of the parameters. Point estimates of elasticities and shadow price ratios undergo sign reversals, and the variances of estimated marginal pdfs become much smaller as more inequality information is incorporated into the estimation process. The estimates obtained from the regularity-constrained models are the only estimates that are theoretically plausible.

REFERENCES

Aigner, D., and S.F. Chu (1968) "On Estimating the Industry Production Function" *American Economic Review* 58:226-39.

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

Albert J.H. and S. Chib (1996) "Computation in Bayesian Econometrics: An Introduction to Markov Chain Monte Carlo" in Hill, R.C. (ed). *Advances in Econometrics Volume 11A: Computational Methods and Applications*. JAI Press, Greenwich. pp. 3-24.

[18] Some of the estimated coverage regions are wide due to the fact that our draws from the marginal pdf of the shadow price ratio become large in absolute value as our draws from the marginal pdf of the denominator in (2.4) get close to zero

[19] Although Geweke (1986) and others argue that sampling theory approaches to imposing these constraints, which can be expressed as inequality constraints on the parameters, are unsatisfactory.

Brummer, B., T. Glauben and G. Thijssen (2002) “Decomposition of Productivity Growth Using Distance Functions: The Case of Dairy Farms in Three European Countries” *American Journal of Agricultural Economics* 84:628-644.

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

Chiang, A.C. (1984) *Fundamental Methods of Mathematical Economics*. 3rd ed. McGraw-Hill. Singapore.

Coelli, T.J. and S. Perelman (1996) “Efficiency Measurement, Multi-output Technologies and Distance Functions: with Application to European Railways”, CREPP Working Paper 96/05, University of Liege.

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.

Coelli, T.J. and Perelman, S. (2000), “Technical Efficiency of European Railways: A Distance Function Approach”, *Applied Economics*, 32, 1967-1976.

Cowie, J. and G. Riddington (1996) “Measuring the Efficiency of European Railways”, Ap*plied Economics*, 28:1027-1035.

Diewert, W.E. and T.J. Wales (1987) “Flexible Forms and Global Curvature Conditions” *Econometrica* 55:43-68.

Dorfman, J.H. and C.S. McIntosh (2001) "Imposing Inequality Restrictions: Efficiency Gains from Economic Theory" *Economics Letters* – forthcoming.

Färe, R. and S. Grosskopf (1994), *Cost and Revenue Constrained Production*, Springer.

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

Färe, R., S. Grosskopf and C.A.K. Lovell (1994) *Production Frontiers*, Cambridge University Press.

Färe, R., S. Grosskopf, C.A.K. Lovell, and S. Yaisawarng (1993) "Derivation of Shadow Prices for Undesirable Outputs: A Distance Function Approach" *Review of Economics and Statistics* 72:374-80.

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.

Fuentes, H.J., E. Grifell-Tatjé and S. Perelman (2001) “A Parametric Distance Function Approach for Malmquist Productivity Index Estimation”, *Journal of Productivity Analysis* 15: 79-94.

Gallant, A.R. and G.H. Golub (1984) "Imposing Curvature Restrictions on Flexible Functional Forms" *Journal of Econometrics* 26:295-321.

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.

Gelfand, A.E. and T-M Lee (1993) "Discussion of the Meeting on the Gibbs Sampler and other Markov Chain Monte Carlo Methods" *Journal of the Royal Statistical Society, Series B* 55:72-73.

Geweke, J. (1986) "Exact Inference in the Inequality Constrained Normal Linear Regression Model" *Journal of Applied Econometrics* 1:127-141.

Griffiths, W.E., O'Donnell, C.J. and A. Tan Cruz (2000) "Imposing regularity conditions on a system of cost and cost-share equations: a Bayesian approach", *Australian Journal of Agricultural and Resource Economics*, 44(1):107-127.

Grosskopf, S., D. Margaritis, and V. Valdmanis (1995) "Estimating Output Substitutability of Hospital Services: A Distance Function Approach." *European Journal of Operational Research* 80:575-87.

Grosskopf, S., K. Hayes and J. Hirschberg (1995) “Fiscal Stress and the Production of Public Safety: Adistance Function Approach” *Journal of Public Economics* 57:277-296.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl and T-C. Lee (1985) *The Theory and Practice of Econometrics*. 2nd ed. New York: John Wiley.

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

Koop, G., J. Osiewalski and M. Steel (1997) "Bayesian Efficiency Analysis Through Individual Effects: Hospital Cost Frontiers" *Journal of Econometrics* 76:77-105.

Koop, G., M. Steel and J. Osiewalski (1995) "Posterior Analysis of Stochastic Frontier Models Using Gibbs Sampling" *Computational Statistics* 10:353-373.

Koop. G. and M. Steel (2001) "Bayesian Analysis of Stochastic Frontier Models" in Baltagi, B. (ed.) *A Companion to Theoretical Econometrics*. MA: Blackwells.

Lau, L.J. (1978) "Testing and Imposing Monoticity[sic], Convexity and Quasi-convexity Constraints" in M. Fuss and D. McFadden (eds.) *Production Economics: A Dual Approach to Theory and Applications*. North Holland. Amsterdam.

Meeusen, W. and J. van den Broeck (1977) "Efficiency Estimation from Cobb-Douglas Production Functions With Composed Error", *International Economic Review*, 18, 435-444.

O'Donnell, C.J., C.R. Shumway and V.E. Ball (1999) "Input demands and inefficiency in U.S. agriculture" *American Journal of Agricultural Economics* 81: 865-880.

Orea, L. (2002) "Parametric Decomposition of a Generalized Malmquist Productivity Index" *Journal of Productivity Analysis* 18:5-22.

Poirier, D. (1995) *Intermediate Statistics and Econometrics: A Comparative Approach* The MIT Press, Cambridge MA.

Rao, A.R. and P. Bhimasankaram (1992) *Linear Algebra.* McGraw Hill. New Delhi.

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.

Roberts, G.O., A. Gelman and W.R. Gilks (1997) "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms" *Annals of Applied Probability* 7:110-120.

Ryan, D.L. and T.J. Wales (2000) "Imposing Local Concavity in the Translog and Generalized Leontief Cost Functions" *Economic Letters* 67:253-260.

Shephard, R.W. (1970), *Theory of Cost and Production Functions*, Princeton, Princeton University Press.

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

Swinton, J.R. (1998) "At What Cost do We Reduce Pollution? Shadow Prices of SO_2 Emissions" The Energy Journal 19:63-83.

Terrell, D. (1996) "Incorporating Monotonicity and Concavity Conditions in Flexible Functional Forms" *Journal of Applied Econometrics* 11:179-194.

van den Broeck, J., G. Koop, J. Osiewalski and M. Steel (1994) "Stochastic Frontier Models: A Bayesian Perspective" *Journal of Econometrics* 46:39-56.

UIC (1988-1993), *International Railway Statistics: Statistics of Individual Railways*, UIC, Paris.

Zellner, A. (1971) *An Introduction to Bayesian Inference in Econometrics*. John Wiley. New York.

DATA APPENDIX

The data were assembled from *International Railways Statistics*, published each year, since1925, by the International Union of Railways (Union International de Chemins de fer, UIC). For each input and output variable discussed below, we indicate the corresponding table containing the annual statistics of individual railways. We also give a summary of the UIC description for each of the selected statistics (UIC, 1988-1993)[20].

Inputs

Staff — *Operating and traffic staff* (Table 31)
Corresponds to the annual mean staff bound to the railway by an employment contract and working in the following activities :
- central and regional operating and traffic departments;
- stations, halts, stopping points, town offices and signaling installations;
- train-accompanying and inspection.

Rolling stock — *Passenger transport stock* (Table 22) and *Freight transport stock* (Table 23)
The available annual mean fleet of coaches multiplied by the average seats and sleep accommodation, and the available annual mean fleet of railway-owned wagons multiplied by the average capacity in tonnes.

Lines — *Lines* (Table 11)
Total length (in km) of lines worked, including electrified and non-electrified lines and broad and narrow gauge lines. Sections permanently out of use are excluded.

Outputs

Passengers — *Revenue-earning passenger traffic* (Table 51)
Number of passenger-kilometers conveyed by rail calculated in accordance with the number of tickets sold multiplied by the kilometric distance for each journey (or by a mean kilometric distance).

Freight — *Freight traffic* (Table 61)
Tonnes-kilometers of revenue-earning traffic carried by rail obtained by multiplying the chargeable weight by the charging distance. This variable includes essentially full wagonloads as well as express parcels and small traffic (including postal packages).

The railway companies included in the data set are:

British Railways (BR); *Swiss Federal Railways* (CFF); *Luxembourg National Railway Company* (CFL); *Hellenic Railways Organisation* (CH); *Irish Transport Company* (CIE); *Portuguese Railways* (CP); *German Federal Railways* (DB); *Danish State Railways* (DSB); *Italian State Railways* (FS); *Netherlands Railways* (NS); *Norwegian State Railways* (NSB); *Austrian Federal Railways* (OBB); *Spanish National Railways* (RENFE); *Swedish State Railways* (SJ); *Belgian National Railway Company* (SNCB/NMBS); *French National Railway Company* (SNCF) and *Finnish State Railways* (VR).

[20] The authors thank Henry-Jean Gathon for making available a computer file containing most of the data used in this study.

Table 1. Intercept Parameters[a]

	Fixed Effects			Random Effects		
	Unconstrained	Monotonicity	Monotonicity & Curvature	Unconstrained	Monotonicity	Monotonicity & Curvature
$\alpha_{0,1}$	0.143 (0.070, 0.217)	0.020 (-0.082, 0.128)	0.035 (-0.084, 0.151)	-	-	-
$\alpha_{0,2}$	0.167 (0.097, 0.237)	0.067 (-0.022, 0.174)	0.041 (-0.049, 0.123)	-	-	-
$\alpha_{0,3}$	0.095 (0.030, 0.158)	0.084 (0.006, 0.196)	-0.013 (-0.094, 0.082)	-	-	-
$\alpha_{0,4}$	0.131 (0.059, 0.208)	-0.027 (-0.130, 0.085)	-0.023 (-0.136, 0.060)	-	-	-
$\alpha_{0,5}$	0.122 (0.052, 0.188)	0.042 (-0.034, 0.136)	-0.009 (-0.093, 0.065)	-	-	-
$\alpha_{0,6}$	0.087 (0.020, 0.156)	-0.011 (-0.098, 0.080)	-0.070 (-0.154, 0.026)	-	-	-
$\alpha_{0,7}$	0.152 (0.072, 0.229)	0.036 (-0.054, 0.126)	0.046 (-0.036, 0.135)	-	-	-
$\alpha_{0,8}$	0.077 (0.005, 0.148)	-0.022 (-0.126, 0.074)	-0.068 (-0.163, 0.021)	-	-	-
$\alpha_{0,9}$	0.049 (-0.022, 0.120)	-0.034 (-0.108, 0.041)	-0.086 (-0.181, 0.003)	-	-	-
$\alpha_{0,10}$	0.099 (0.027, 0.172)	0.014 (-0.065, 0.087)	0.057 (-0.045, 0.139)	-	-	-
$\alpha_{0,11}$	0.069 (0.002, 0.137)	-0.042 (-0.115, 0.035)	-0.087 (-0.172, 0.002)	-	-	-
$\alpha_{0,12}$	0.055 (-0.018, 0.128)	-0.030 (-0.121, 0.052)	-0.031 (-0.112, 0.052)	-	-	-
$\alpha_{0,13}$	0.059 (-0.016, 0.135)	-0.042 (-0.141, 0.040)	-0.093 (-0.211, -0.003)	-	-	-
$\alpha_{0,14}$	0.080 (0.008, 0.151)	-0.015 (-0.115, 0.094)	-0.089 (-0.225, 0.007)	-	-	-
$\alpha_{0,15}$	0.067 (-0.012, 0.145)	0.005 (-0.082, 0.106)	-0.021 (-0.111, 0.065)	-	-	-
$\alpha_{0,16}$	0.047 (-0.021, 0.111)	-0.078 (-0.163, 0.017)	-0.134 (-0.237, -0.050	-	-	-
$\alpha_{0,17}$	0.081 (0.009, 0.153)	-0.016 (-0.081, 0.059)	-0.012 (-0.092, 0.073)	-	-	-
a_0	-	-	-	0.054 (0.005, 0.102)	-0.029 (-0.069, 0.011)	-0.078 (-0.110, -0.043)

[a] Numbers in parentheses are 95% posterior coverage regions.

Table 2. Slope Parameters[a]

	Fixed Effects			Random Effects		
	Unconstrained	Monotonicity	Monotonicity & Curvature	Unconstrained	Monotonicity	Monotonicity & Curvature
a_1	0.583 (0.536, 0.630)	0.519 (0.473, 0.568)	0.536 (0.483, 0.577)	0.571 (0.526, 0.615)	0.511 (0.463, 0.558)	0.518 (0.471, 0.563)
a_{11}	0.361 (0.260, 0.465)	0.113 (0.014, 0.196)	0.255 (0.250, 0.265)	0.372 (0.276, 0.469)	0.103 (0.010, 0.186)	0.256 (0.250, 0.266)
b_1	-0.662 (-0.767, -0.557)	-0.509 (-0.590, -0.454)	-0.439 (-0.488, -0.397)	-0.632 (-0.729, -0.533)	-0.491 (-0.543, -0.449)	-0.438 (-0.486, -0.399)
b_2	0.017 (-0.086, 0.124)	-0.140 (-0.213, -0.066)	-0.184 (-0.236, -0.134)	-0.013 (-0.113, 0.086)	-0.175 (-0.240, -0.106)	-0.193 (-0.249, -0.134)
b_3	-0.477 (-0.524, -0.435)	-0.428 (-0.493, -0.374)	-0.428 (-0.485, -0.371)	-0.473 (-0.515, -0.431)	-0.409 (-0.465, -0.352)	-0.415 (-0.468, -0.364)
b_{11}	0.783 (0.193, 1.375)	0.237 (0.093, 0.379)	0.268 (0.161, 0.365)	0.736 (0.197, 1.276)	0.222 (0.073, 0.358)	0.282 (0.179, 0.390)
b_{12}	-1.100 (-1.609, -0.614)	-0.043 (-0.115, 0.026)	-0.073 (-0.151, 0.009)	-1.025 (-1.482, -0.569)	-0.050 (-0.117, 0.013)	-0.084 (-0.170, 0.003)
b_{13}	0.663 (0.447, 0.881)	0.026 (-0.073, 0.144)	-0.046 (-0.099, 0.006)	0.628 (0.422, 0.832)	0.043 (-0.070, 0.189)	-0.045 (-0.091, 0.004)
b_{22}	1.188 (0.717, 1.690)	0.009 (-0.060, 0.101)	0.118 (0.023, 0.212)	1.094 (0.645, 1.544)	0.001 (-0.062, 0.078)	0.130 (0.040, 0.226)
b_{23}	-0.199 (-0.348, -0.048)	0.039 (-0.046, 0.124)	-0.029 (-0.083, 0.032)	-0.163 (-0.302, -0.023)	0.048 (-0.044, 0.149)	-0.025 (-0.075, 0.026)
b_{33}	-0.706 (-0.911, -0.504)	-0.197 (-0.294, -0.082)	-0.020 (-0.081, 0.038)	-0.714 (-0.901, -0.526)	-0.208 (-0.321, -0.082)	-0.022 (-0.081, 0.030)
g_{11}	-0.679 (-0.802, -0.563)	-0.284 (-0.358, -0.201)	-0.325 (-0.378, -0.262)	-0.687 (-0.802, -0.572)	-0.306 (-0.384, -0.232)	-0.344 (-0.404, -0.281)
g_{21}	0.565 (0.419, 0.718)	0.078 (0.016, 0.139)	0.153 (0.084, 0.207)	0.550 (0.409, 0.690)	0.098 (0.035, 0.165)	0.165 (0.103, 0.223)
g_{31}	0.148 (0.049, 0.250)	0.163 (0.117, 0.206)	0.170 (0.131, 0.205)	0.171 (0.076, 0.267)	0.162 (0.105, 0.209)	0.170 (0.132, 0.202)

[a] Numbers in parentheses are 95% posterior coverage regions.

Table 3. Technical Efficiency Scores (i = 1, ..., 17)[a].

	Fixed Effects			Random Effects		
Firm	Unconstrained	Monotonicity	Monotonicity & Curvature	Unconstrained	Monotonicity	Monotonicity & Curvature
1	0.958 (0.902, 1.000)	0.904 (0.798, 1.000)	0.932 (0.829, 1.000)	0.942 (0.889, 0.990)	0.951 (0.885, 0.996)	0.948 (0.874, 0.996)
2	0.982 (0.931, 1.000)	0.947 (0.855, 1.000)	0.938 (0.843, 1.000)	0.924 (0.870, 0.979)	0.936 (0.863, 0.993)	0.937 (0.857, 0.994)
3	0.914 (0.852, 0.978)	0.962 (0.886, 1.000)	0.888 (0.818, 0.986)	0.968 (0.925, 0.997)	0.948 (0.880, 0.995)	0.955 (0.889, 0.997)
4	0.948 (0.893, 1.000)	0.863 (0.750, 0.968)	0.879 (0.784, 0.964)	0.954 (0.902, 0.995)	0.971 (0.921, 0.998)	0.965 (0.906, 0.998)
5	0.939 (0.881, 1.000)	0.924 (0.825, 1.000)	0.892 (0.801, 0.975)	0.955 (0.906, 0.995)	0.960 (0.902, 0.997)	0.958 (0.893, 0.997)
6	0.907 (0.848, 0.971)	0.876 (0.790, 0.964)	0.839 (0.768, 0.934)	0.970 (0.929, 0.997)	0.969 (0.918, 0.998)	0.973 (0.925, 0.998)
7	0.968 (0.909, 1.000)	0.919 (0.818, 1.000)	0.942 (0.860, 1.000)	0.933 (0.878, 0.985)	0.943 (0.871, 0.995)	0.936 (0.854, 0.994)
8	0.898 (0.840, 0.956)	0.867 (0.777, 0.950)	0.841 (0.751, 0.921)	0.977 (0.941, 0.998)	0.974 (0.929, 0.999)	0.976 (0.931, 0.999)
9	0.873 (0.815, 0.932)	0.856 (0.784, 0.941)	0.826 (0.744, 0.909)	0.984 (0.956, 0.999)	0.977 (0.934, 0.999)	0.977 (0.935, 0.999)
10	0.917 (0.862, 0.973)	0.898 (0.820, 0.984)	0.952 (0.867, 1.000)	0.970 (0.928, 0.998)	0.959 (0.898, 0.997)	0.948 (0.874, 0.996)
11	0.891 (0.834, 0.946)	0.850 (0.758, 0.933)	0.825 (0.741, 0.911)	0.979 (0.945, 0.999)	0.977 (0.935, 0.999)	0.978 (0.937, 0.999)
12	0.878 (0.819, 0.935)	0.859 (0.778, 0.933)	0.872 (0.806, 0.952)	0.983 (0.952, 0.999)	0.971 (0.923, 0.998)	0.968 (0.914, 0.998)
13	0.882 (0.827, 0.940)	0.850 (0.764, 0.932)	0.820 (0.725, 0.906)	0.982 (0.952, 0.999)	0.978 (0.939, 0.999)	0.976 (0.932, 0.999)
14	0.900 (0.838, 0.960)	0.873 (0.782, 1.000)	0.823 (0.745, 0.904)	0.976 (0.939, 0.998)	0.972 (0.924, 0.998)	0.973 (0.925, 0.998)
15	0.889 (0.830, 0.948)	0.890 (0.798, 0.989)	0.881 (0.786, 0.971)	0.979 (0.944, 0.999)	0.967 (0.914, 0.998)	0.962 (0.900, 0.998)
16	0.871 (0.815, 0.927)	0.819 (0.736, 0.917)	0.787 (0.687, 0.858)	0.985 (0.957, 0.999)	0.982 (0.949, 0.999)	0.979 (0.941, 0.999)
17	0.901 (0.840, 0.964)	0.872 (0.777, 0.968)	0.889 (0.816, 0.979)	0.975 (0.937, 0.998)	0.967 (0.913, 0.998)	0.972 (0.924, 0.998)

[a] Numbers in parentheses are 95% posterior coverage regions.

Table 4. Elasticities for Input 1 ($t = 6$; $i = 1, ..., 17$)[a].

	Fixed Effects			Random Effects		
Firm	Unconstrained	Monotonicity	Monotonicity & Curvature	Unconstrained	Monotonicity	Monotonicity & Curvature
1	-0.234 (-0.467, 0.004)	-0.540 (-0.628, -0.451)	-0.522 (-0.603, -0.446)	-0.256 (-0.465, -0.045)	-0.533 (-0.613, -0.448)	-0.525 (-0.600, -0.448)
2	-1.242 (-1.474, -1.017)	-0.662 (-0.795, -0.555)	-0.518 (-0.581, -0.458)	-1.201 (-1.417, -0.986)	-0.662 (-0.790, -0.552)	-0.522 (-0.581, -0.462)
3	-1.640 (-2.069, -1.222)	-0.924 (-1.187, -0.665)	-0.576 (-0.777, -0.332)	-1.540 (-1.921, -1.160)	-0.877 (-1.143, -0.595)	-0.571 (-0.783, -0.370)
4	-1.842 (-2.220, -1.466)	-1.248 (-1.366, -1.120)	-1.139 (-1.230, -1.031)	-1.812 (-2.155, -1.471)	-1.231 (-1.341, -1.106)	-1.169 (-1.248, -1.080)
5	-0.244 (-0.562, 0.106)	-1.130 (-1.259, -1.002)	-0.949 (-1.081, -0.821)	-0.303 (-0.617, 0.011)	-1.094 (-1.217, -0.969)	-0.958 (-1.076, -0.847)
6	-0.676 (-0.937, -0.401)	-1.018 (-1.093, -0.938)	-0.887 (-0.958, -0.811)	-0.716 (-0.964, -0.468)	-1.007 (-1.074, -0.932)	-0.900 (-0.968, -0.821)
7	-0.500 (-0.722, -0.287)	-0.176 (-0.287, -0.111)	-0.181 (-0.279, -0.121)	-0.448 (-0.650, -0.243)	-0.164 (-0.248, -0.110)	-0.173 (-0.252, -0.125)
8	-0.782 (-1.039, -0.518)	-0.964 (-1.048, -0.878)	-0.805 (-0.877, -0.728)	-0.804 (-1.047, -0.560)	-0.953 (-1.023, -0.870)	-0.815 (-0.885, -0.737)
9	-1.010 (-1.200, -0.827)	-0.523 (-0.620, -0.438)	-0.522 (-0.619, -0.443)	-0.985 (-1.160, -0.810)	-0.532 (-0.620, -0.452)	-0.531 (-0.615, -0.451)
10	-1.342 (-1.625, -1.059)	-1.077 (-1.164, -0.979)	-0.961 (-1.031, -0.883)	-1.364 (-1.625, -1.105)	-1.092 (-1.170, -0.997)	-0.984 (-1.056, -0.896)
11	-0.134 (-0.369, 0.105)	-0.797 (-0.963, -0.632)	-0.653 (-0.801, -0.521)	-0.137 (-0.356, 0.082)	-0.740 (-0.894, -0.585)	-0.650 (-0.780, -0.535)
12	-0.439 (-0.658, -0.228)	-0.397 (-0.539, -0.293)	-0.271 (-0.341, -0.209)	-0.408 (-0.608, -0.209)	-0.378 (-0.490, -0.272)	-0.260 (-0.332, -0.191)
13	-0.866 (-1.128, -0.601)	-0.741 (-0.839, -0.628)	-0.723 (-0.781, -0.663)	-0.849 (-1.095, -0.605)	-0.720 (-0.826, -0.589)	-0.736 (-0.789, -0.678)
14	0.383 (0.076, 0.692)	-0.362 (-0.556, -0.177)	-0.260 (-0.416, -0.126)	0.409 (0.125, 0.692)	-0.291 (-0.466, -0.118)	-0.240 (-0.385, -0.110)
15	-0.949 (-1.142, -0.758)	-0.521 (-0.663, -0.408)	-0.370 (-0.439, -0.301)	-0.894 (-1.073, -0.714)	-0.504 (-0.627, -0.386)	-0.364 (-0.436, -0.295)
16	-0.433 (-0.632, -0.232)	-0.287 (-0.388, -0.198)	-0.317 (-0.414, -0.251)	-0.402 (-0.593, -0.207)	-0.275 (-0.368, -0.184)	-0.316 (-0.393, -0.256)
17	0.080 (-0.206, 0.368)	-0.439 (-0.629, -0.256)	-0.297 (-0.457, -0.150)	0.119 (-0.144, 0.381)	-0.371 (-0.539, -0.202)	-0.279 (-0.432, -0.141)

[a] Numbers in parentheses are 95% posterior coverage regions.

Table 5. Elasticities for Input 2 (t = 6; i = 1, ..., 17)[a].

	Fixed Effects			Random Effects		
Firm	Unconstrained	Monotonicity	Monotonicity & Curvature	Unconstrained	Monotonicity	Monotonicity & Curvature
1	-0.357 (-0.573, -0.145)	-0.092 (-0.174, -0.022)	-0.150 (-0.225, -0.089)	-0.333 (-0.520, -0.146)	-0.113 (-0.201, -0.033)	-0.155 (-0.223, -0.095)
2	0.094 (-0.061, 0.250)	-0.153 (-0.257, -0.048)	-0.147 (-0.214, -0.086)	0.047 (-0.098, 0.191)	-0.187 (-0.299, -0.090)	-0.160 (-0.227, -0.093)
3	0.239 (-0.169, 0.664)	-0.249 (-0.483, -0.032)	-0.343 (-0.535, -0.139)	0.127 (-0.243, 0.499)	-0.292 (-0.566, -0.055)	-0.378 (-0.553, -0.200)
4	0.981 (0.615, 1.362)	-0.037 (-0.094, -0.007)	-0.050 (-0.124, -0.013)	0.905 (0.563, 1.245)	-0.035 (-0.084, -0.008)	-0.051 (-0.120, -0.014)
5	-0.729 (-1.032, -0.444)	-0.083 (-0.191, -0.011)	-0.270 (-0.376, -0.132)	-0.684 (-0.957, -0.409)	-0.077 (-0.177, -0.009)	-0.293 (-0.404, -0.191)
6	-0.371 (-0.598, -0.152)	-0.072 (-0.142, -0.026)	-0.162 (-0.236, -0.089)	-0.348 (-0.553, -0.142)	-0.071 (-0.129, -0.032)	-0.176 (-0.258, -0.113)
7	0.136 (-0.052, 0.334)	-0.158 (-0.268, -0.058)	-0.164 (-0.245, -0.094)	0.096 (-0.093, 0.283)	-0.210 (-0.314, -0.105)	-0.167 (-0.251, -0.089)
8	-0.360 (-0.562, -0.161)	-0.100 (-0.164, -0.042)	-0.182 (-0.247, -0.105)	-0.351 (-0.541, -0.162)	-0.107 (-0.170, -0.055)	-0.198 (-0.275, -0.134)
9	0.332 (0.154, 0.515)	-0.091 (-0.176, -0.019)	-0.052 (-0.131, -0.005)	0.299 (0.127, 0.471)	-0.119 (-0.211, -0.032)	-0.049 (-0.115, -0.004)
10	0.061 (-0.147, 0.269)	-0.050 (-0.133, -0.004)	-0.044 (-0.108, -0.004)	0.058 (-0.136, 0.252)	-0.048 (-0.124, -0.004)	-0.050 (-0.129, -0.004)
11	-0.388 (-0.527, -0.246)	-0.136 (-0.233, -0.048)	-0.328 (-0.413, -0.190)	-0.389 (-0.524, -0.254)	-0.156 (-0.254, -0.065)	-0.350 (-0.443, -0.252)
12	-0.407 (-0.546, -0.273)	-0.191 (-0.292, -0.078)	-0.257 (-0.334, -0.174)	-0.425 (-0.555, -0.294)	-0.236 (-0.341, -0.136)	-0.276 (-0.358, -0.186)
13	0.438 (0.209, 0.672)	-0.069 (-0.136, -0.013)	-0.110 (-0.165, -0.063)	0.403 (0.185, 0.620)	-0.086 (-0.162, -0.017)	-0.110 (-0.161, -0.065)
14	-0.518 (-0.739, -0.300)	-0.196 (-0.326, -0.063)	-0.409 (-0.518, -0.257)	-0.528 (-0.739, -0.318)	-0.242 (-0.367, -0.109)	-0.431 (-0.550, -0.301)
15	-0.014 (-0.161, 0.140)	-0.189 (-0.306, -0.067)	-0.224 (-0.309, -0.142)	-0.066 (-0.205, 0.071)	-0.233 (-0.359, -0.122)	-0.241 (-0.326, -0.151)
16	0.148 (-0.042, 0.340)	-0.120 (-0.220, -0.037)	-0.137 (-0.216, -0.073)	0.123 (-0.062, 0.305)	-0.160 (-0.269, -0.060)	-0.136 (-0.209, -0.073)
17	-0.409 (-0.625, -0.197)	-0.207 (-0.334, -0.072)	-0.403 (-0.516, -0.246)	-0.436 (-0.639, -0.232)	-0.253 (-0.386, -0.122)	-0.427 (-0.549, -0.293)

[a] Numbers in parentheses are 95% posterior coverage regions.

Table 6. Elasticities for Input 3 (t = 6; i = 1, ..., 17)[a].

	Fixed Effects			Random Effects		
Firm	Unconstrained	Monotonicity	Monotonicity & Curvature	Unconstrained	Monotonicity	Monotonicity & Curvature
1	-0.480 (-0.589, -0.372)	-0.439 (-0.510, -0.353)	-0.362 (-0.442, -0.288)	-0.485 (-0.586, -0.384)	-0.420 (-0.494, -0.331)	-0.350 (-0.426, -0.277)
2	0.267 (0.113, 0.425)	-0.190 (-0.285, -0.116)	-0.329 (-0.380, -0.278)	0.277 (0.131, 0.423)	-0.168 (-0.264, -0.076)	-0.317 (-0.364, -0.273)
3	0.317 (0.109, 0.524)	-0.104 (-0.238, -0.011)	-0.311 (-0.431, -0.206)	0.331 (0.133, 0.526)	-0.121 (-0.270, -0.015)	-0.305 (-0.399, -0.205)
4	-0.430 (-0.670, -0.191)	-0.133 (-0.257, -0.031)	-0.079 (-0.146, -0.028)	-0.387 (-0.609, -0.162)	-0.151 (-0.305, -0.034)	-0.073 (-0.144, -0.024)
5	-0.200 (-0.393, -0.015)	-0.223 (-0.340, -0.085)	-0.092 (-0.197, -0.014)	-0.211 (-0.388, -0.034)	-0.256 (-0.390, -0.122)	-0.091 (-0.185, -0.016)
6	0.024 (-0.096, 0.145)	-0.173 (-0.233, -0.106)	-0.121 (-0.187, -0.067)	0.024 (-0.088, 0.137)	-0.183 (-0.248, -0.116)	-0.116 (-0.183, -0.065)
7	-0.774 (-0.859, -0.696)	-0.592 (-0.722, -0.492)	-0.598 (-0.692, -0.498)	-0.770 (-0.851, -0.689)	-0.552 (-0.661, -0.444)	-0.580 (-0.670, -0.492)
8	0.167 (0.078, 0.257)	-0.154 (-0.199, -0.113)	-0.157 (-0.211, -0.118)	0.167 (0.083, 0.252)	-0.160 (-0.201, -0.124)	-0.151 (-0.199, -0.115)
9	-0.333 (-0.416, -0.253)	-0.361 (-0.466, -0.282)	-0.380 (-0.457, -0.303)	-0.316 (-0.393, -0.241)	-0.327 (-0.409, -0.248)	-0.365 (-0.440, -0.292)
10	0.426 (0.292, 0.558)	-0.023 (-0.068, -0.001)	-0.065 (-0.119, -0.020)	0.443 (0.316, 0.569)	-0.018 (-0.052, -0.001)	-0.057 (-0.111, -0.015)
11	-0.879 (-1.078, -0.687)	-0.474 (-0.589, -0.355)	-0.328 (-0.401, -0.265)	-0.886 (-1.064, -0.707)	-0.499 (-0.626, -0.374)	-0.323 (-0.391, -0.264)
12	-0.140 (-0.246, -0.027)	-0.406 (-0.488, -0.332)	-0.477 (-0.533, -0.413)	-0.150 (-0.256, -0.044)	-0.383 (-0.467, -0.289)	-0.465 (-0.518, -0.412)
13	-0.885 (-1.086, -0.692)	-0.431 (-0.536, -0.326)	-0.319 (-0.394, -0.251)	-0.865 (-1.043, -0.685)	-0.427 (-0.546, -0.310)	-0.308 (-0.383, -0.241)
14	-1.372 (-1.586, -1.153)	-0.738 (-0.855, -0.617)	-0.580 (-0.661, -0.501)	-1.390 (-1.586, -1.196)	-0.747 (-0.865, -0.619)	-0.570 (-0.639, -0.501)
15	-0.031 (-0.152, 0.096)	-0.317 (-0.405, -0.249)	-0.432 (-0.487, -0.372)	-0.028 (-0.145, 0.089)	-0.298 (-0.386, -0.209)	-0.420 (-0.472, -0.369)
16	-0.912 (-1.026, -0.806)	-0.591 (-0.704, -0.484)	-0.532 (-0.629, -0.438)	-0.906 (-1.008, -0.803)	-0.558 (-0.659, -0.451)	-0.516 (-0.604, -0.427)
17	-1.113 (-1.265, -0.957)	-0.644 (-0.734, -0.554)	-0.544 (-0.611, -0.478)	-1.127 (-1.266, -0.990)	-0.654 (-0.743, -0.564)	-0.536 (-0.594, -0.478)

[a] Numbers in parentheses are 95% posterior coverage regions.

Table 7. Output Shadow Price Ratio (t = 6; i = 1, ..., 17)[a].

Firm	Fixed Effects			Random Effects		
	Unconstrained	Monotonicity	Monotonicity & Curvature	Unconstrained	Monotonicity	Monotonicity & Curvature
1	0.644	0.656	0.926	0.653	0.594	0.822
	(0.407, 0.963)	(0.471, 0.859)	(0.687, 1.182)	(0.427, 0.947)	(0.409, 0.809)	(0.632, 1.067)
2	0.553	0.519	0.572	0.499	0.489	0.532
	(0.370, 0.772)	(0.406, 0.659)	(0.455, 0.704)	(0.342, 0.683)	(0.378, 0.617)	(0.420, 0.664)
3	0.996	3.020	1.028	0.791	3.292	1.077
	(0.002, 2.449)	(1.670, 4.917)	(0.698, 1.434)	(-0.050, 1.951)	(1.677, 5.869)	(0.788, 1.421)
4	-2.039	4.305	17.673	-3.130	3.692	18.766
	(-9.048, -1.543)	(1.050, 13.575)	(8.533, 27.143)	(-7.219, -1.551)	(1.064, 10.374)	(9.178, 28.984
5	0.231	1.729	1.215	0.261	1.464	1.143
	(0.060, 0.456)	(0.980, 2.841)	(0.857, 1.820)	(0.091, 0.492)	(0.903, 2.315)	(0.803, 1.630)
6	0.288	0.702	0.826	0.305	0.606	0.758
	(0.151, 0.465)	(0.434, 1.033)	(0.637, 1.143)	(0.174, 0.476)	(0.406, 0.861)	(0.568, 1.014)
7	3.050	1.029	1.272	2.710	1.025	1.155
	(1.840, 4.945)	(0.727, 1.449)	(0.843, 1.685)	(1.680, 4.210)	(0.706, 1.411)	(0.796, 1.534)
8	0.251	0.666	0.687	0.256	0.590	0.639
	(0.145, 0.377)	(0.452, 0.916)	(0.550, 0.893)	(0.159, 0.373)	(0.432, 0.779)	(0.496, 0.819)
9	2.170	0.486	0.924	2.087	0.455	0.826
	(1.007, 4.768)	(0.332, 0.656)	(0.654, 1.225)	(0.994, 4.039)	(0.290, 0.655)	(0.607, 1.097)
10	0.304	0.312	0.567	0.307	0.264	0.510
	(0.161, 0.516)	(0.168, 0.490)	(0.420, 0.799)	(0.177, 0.497)	(0.152, 0.403)	(0.370, 0.701)
11	1.377	4.595	2.477	1.448	4.506	2.428
	(0.915, 1.947)	(3.251, 6.493)	(2.119, 2.868)	(0.988, 2.025)	(3.214, 6.306)	(2.077, 2.789)
12	0.441	0.887	0.678	0.384	0.851	0.614
	(0.243, 0.678)	(0.619, 1.278)	(0.524, 0.840)	(0.208, 0.583)	(0.611, 1.140)	(0.467, 0.778)
13	14.276	1.750	3.041	-24.761	1.700	2.774
	(-69.349, 71.487)	(1.160, 2.477)	(2.101, 4.015)	(-76.863, 74.817)	(1.042, 2.562)	(2.058, 3.541)
14	2.836	7.666	3.295	2.809	7.968	3.208
	(1.696, 4.270)	(4.404, 12.978)	(2.747, 3.804)	(1.777, 4.089)	(5.025, 12.574)	(2.667, 3.672)
15	0.822	0.962	0.775	0.724	0.947	0.727
	(0.492, 1.217)	(0.727, 1.270)	(0.628, 0.942)	(0.446, 1.054)	(0.717, 1.230)	(0.584, 0.886)
16	5.168	1.058	1.536	3.843	1.037	1.385
	(2.118, 7.522)	(0.763, 1.431)	(1.037, 2.042)	(2.037, 6.927)	(0.700, 1.454)	(0.976, 1.827)
17	2.406	6.649	2.770	2.316	6.974	2.726
	(1.369, 3.697)	(3.947, 10.961)	(2.426, 3.087)	(1.400, 3.438)	(4.410, 10.903)	(2.371, 3.020)

[a] Numbers in parentheses are 95% posterior coverage regions.

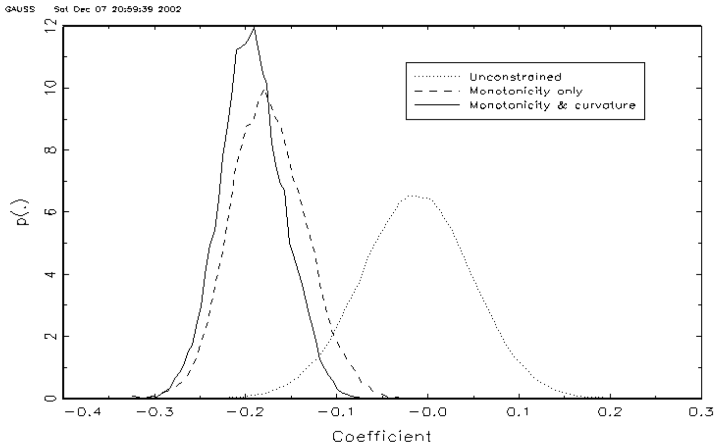

Fig. 1. Random Effects Estimates of the Posterior Pdf for b_2

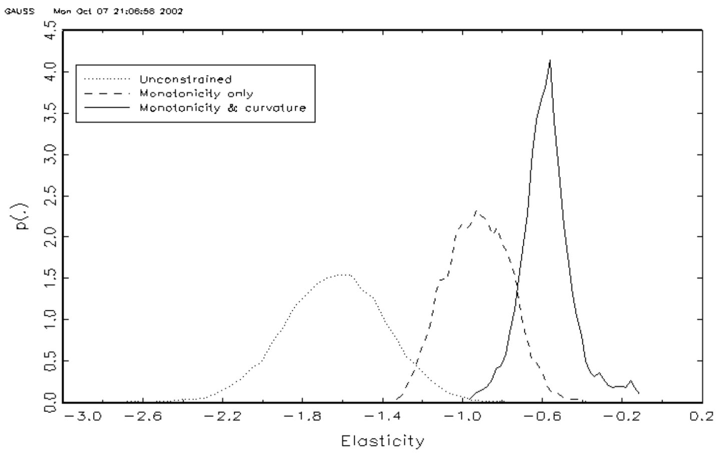

Fig. 2. Fixed Effects Estimates of Elasticity for Input 1, Firm i = 3, Period t = 6

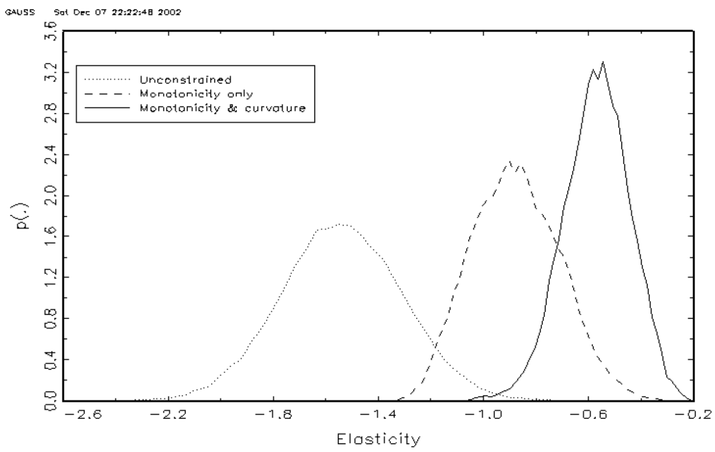

Fig. 3. Random Effects Estimates of Elasticity for Input 1, Firm i = 3, Period t = 6

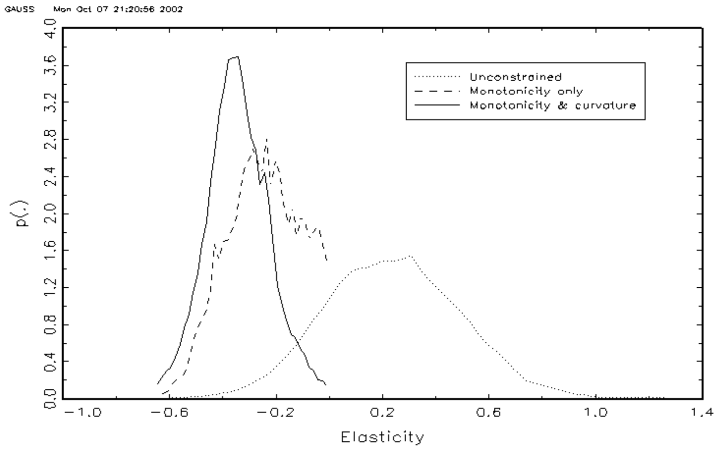

Fig. 4. Fixed Effects Estimates of Elasticity for Input 2 Firm *i* = 3, Period *t* = 6

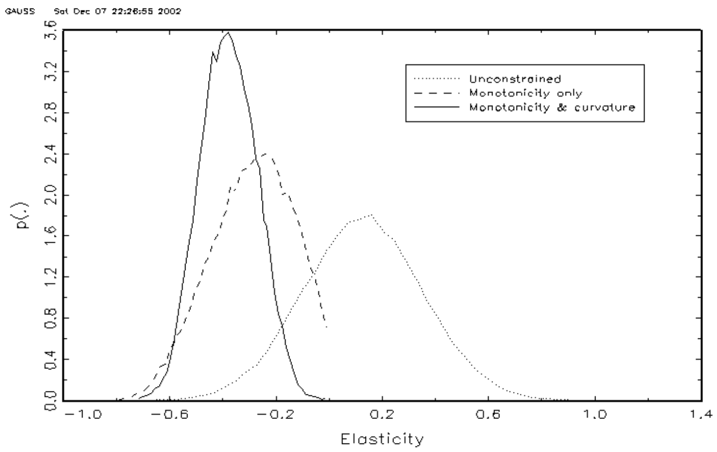

Fig. 5. Random Effects Estimates of Elasticity for Input 2 Firm *i* = 3, Period *t* = 6

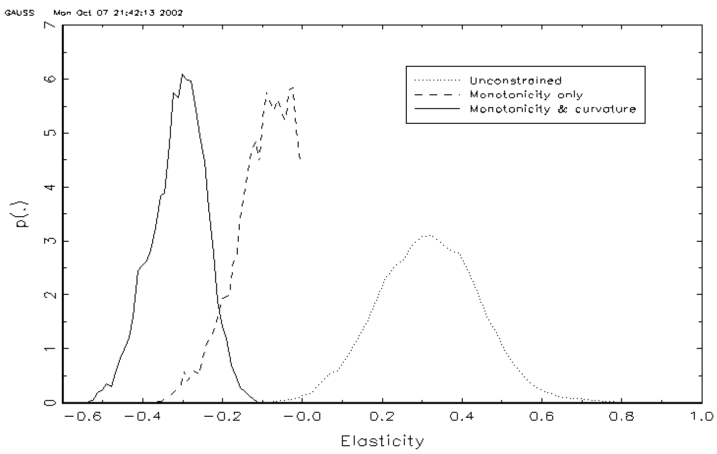

Fig. 6. Fixed Effects Estimates of Elasticity for Input 3 Firm *i* = 3, Period *t* = 6

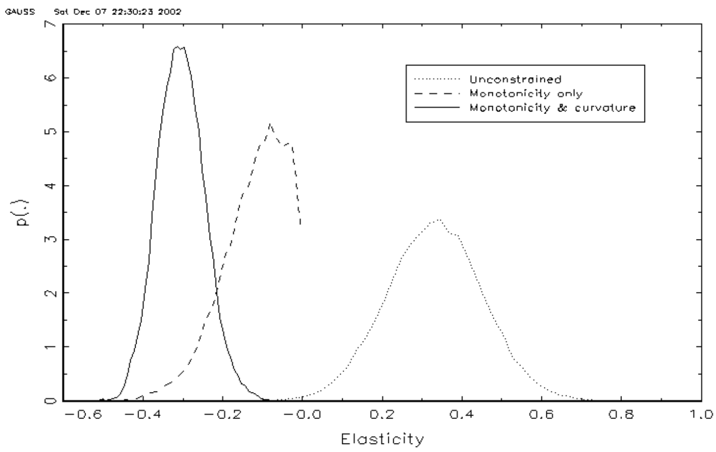

Fig. 7. Random Effects Estimates of Elasticity for Input 3 Firm *i* = 3, Period *t* = 6

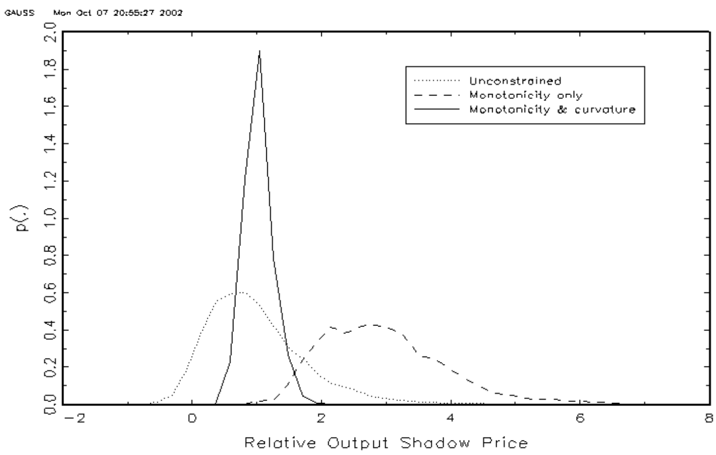

Fig. 8. Fixed Effects Estimates of Output Shadow Price Ratio, Firm *i* = 3, Period *t* = 6

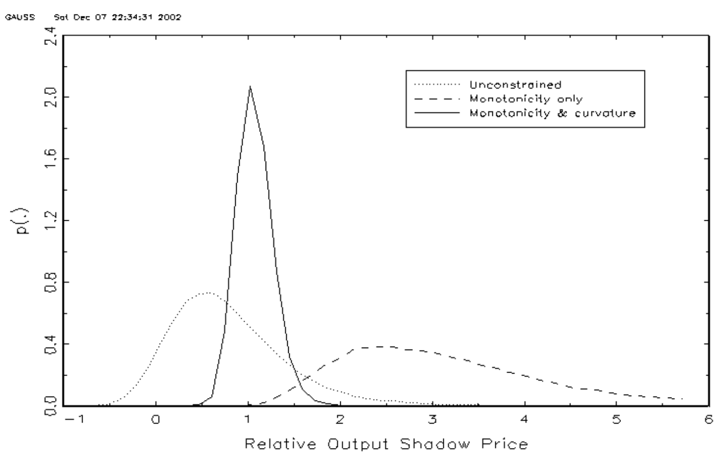

Fig. 9. Random Effects Estimates of Output Shadow Price Ratio, Firm *i* = 3, Period *t* = 6