Bayesian Analysis of Input-Oriented Technical Efficiency

Subal C. Kumbhakar
Department of Economics
State University of New York
Binghamton, NY 13902, USA.
Phone: (607) 777 4762, Fax: (607) 777 2681
E-mail: kkar@binghamton.edu

and

Efthymios G. Tsionas
Department of Economics
Athens University of Economics and Business
76 Patission Street, 104 34 Athens, Greece.
Phone: (3021) 0820 3388, Fax: (3021) 0820 3301
E-mail: tsionas@aueb.gr

September 8, 2003

Abstract

This paper deals with Bayesian estimation of input-oriented (IO) technical efficiency using a stochastic production frontier approach. Although the IO technical efficiency concept is commonly used in theoretical works, it is never estimated in practice using a production function. We provide inferences for parameters and efficiency using Bayesian methods based on Markov Chain Monte Carlo techniques, especially the Gibbs sampler with data augmentation. Both cross-sectional and panel data models are developed. To emphasize the point that estimated efficiency, returns to scale, *etc*., might differ depending on whether one specifies an IO or output-oriented (OO) technical efficiency term within the context of the model, we compare results from the IO and OO models using different priors. The proposed techniques are illustrated using dairy data.

JEL Classification No.: C11, C21, C23

Keywords: Input and output technical efficiency, translog production function, returns to scale, and technical change.

1. Introduction

Two measures of technical efficiency are primarily employed in the efficiency literature. These are input- and output-oriented (henceforth IO and OO) technical inefficiency.[1] The stochastic frontier (SF) model developed by Aigner et al. (1977) (ALS) and Meeusen and van den Broeck (1977) used the OO measure of technical inefficiency. Since then the entire literature on SF models has been over run with models representing OO technical inefficiency (see Chapter 3 of Kumbhakar and Lovell (2000)). Although there are some basic differences between the IO and OO models and some of these differences are well known, no one has estimated a SF model econometrically using IO technical inefficiency.[2]

In this paper we use a Bayesian approach to estimate production function models incorporating IO technical efficiency. Bayesian analysis of a stochastic frontier function (using OO technical inefficiency) was first proposed by van den Broeck, Koop, Osiewalski, and Steel (1994). Koop, Steel, and Osiewalski (1995) first used the Gibbs sampler as a numerical technique where it was shown that Gibbs sampling has a great advantage over importance sampling. Koop, Osiewalski, and Steel (1997) proposed measuring OO technical inefficiency in panel data models where technical inefficiency is time-invariant. Our model is fundamentally different from Koop et al (1997) on economic as well as technical grounds. From the economic point of view we show that the choice of orientation has substantive implications not only for efficiency measurement but also for the measurement of input elasticities, returns to scale and technical change. From the technical point of view, the OO model is simply a shift of the frontier whereas the IO model is a "twist" of the frontier. The econometric implication is that instead of a location mixture of normal errors in the OO model, in the IO model we have a nonlinear error-components model, the estimation of which requires specialized techniques.

We provide inferences for parameters and efficiency using Markov Chain Monte Carlo techniques, especially the Gibbs sampler with data augmentation. Both cross-sectional and panel data models are developed. We compare results from the IO and OO models using different priors to emphasize the point that estimated efficiency, returns to scale, etc., might differ

[1] See Fare and Lovell (1978) for an earlier discussion on these issues.
[2] On the contrary the IO model has been estimated by many using the DEA approach (see Ray (2003) and the references cited therein).

depending on whether one uses the IO or the OO model. The techniques proposed in the paper are illustrated using Spanish dairy farm data.

The rest of the paper is organized as follows. Section 2 introduces the IO and OO models. The econometric models are discussed in Section 3. Section 4 presents some results on the issues mentioned above using a sample of Spanish dairy farm data. Finally, Section 5 concludes the paper.

2. The IO and OO models

Our approach to the efficiency measurement problem is parametric. We start with a single output production technology where Y is a scalar output and X is a vector of inputs. Then the production technology with the IO measure of technical inefficiency can be expressed as

$$Y_i = f(X_i \cdot \Theta_i)\,,\; i = 1, \ldots, n, \tag{1}$$

where Y_i is a scalar output, X_i is the $J \times 1$ vector of inputs used, $\Theta_i = X^e_{ji} / X_{ji} \leq 1 \; \forall \; j$ is input-oriented efficiency, and $X^e_i = \Theta_i X_i$ is the input vector in efficiency units. Finally, the subscript i indexes firms. The IO technical inefficiency for firm i is defined as $\ln \Theta_i \leq 0$ and is interpreted as the rate at which all the inputs can be reduced without reducing output.

On the other hand, the technology with the OO measure of technical inefficiency is modeled as

$$Y_i = f(X_i) \cdot \Lambda_i \tag{2}$$

where $\Lambda_i = Y_i / f(X_i) \leq 1$ represents output-oriented efficiency. Since it is the ratio of actual output to the frontier (maximum possible) output, $\Lambda_i \leq 1$ and therefore, $\ln \Lambda_i \leq 0$, which is the measure of OO technical inefficiency. It shows the percent by which actual output could be increased without increasing inputs. It can also be viewed as output shortfall (percent by which the actual output falls short of the maximum possible output), given the inputs.

Before going further, it is perhaps important to ask the following questions. Why do we need to estimate a model with IO technical inefficiency? Can one translate the IO model into an OO model and use the techniques developed for OO models? The answer depends on the type of production function one assumes to represent the underlying technology. For example, that if $f(\cdot)$ is homogeneous of degree *r* then $\Theta_i^{\,r} = \Lambda_i$ which holds irrespective of the values of *X* and *Y*. In such a case it is not necessary to deal with the IO model separately. This is because the econometric model corresponding to the IO and OO models will be the same. Consequently, information regarding IO technical efficiency can be obtained from estimating the OO model. However, if the production function is not homogeneous, the econometric model corresponding to the IO and OO models will be different. That is, econometrically the IO model cannot be estimated using the same tools that are used to estimate the OO model. And this is what we do in this paper.

3. The econometric model and Bayesian inference procedures

3.1. Econometric model

Consider the IO model in (1). Let lower case letters indicate the log of a variable, and assume that $f(.)$ has a translog form. With these in place the production function in (1) can be written as

$$\begin{aligned} y_i &= \beta_0 + (x_i - \theta_i 1_J)'\beta + \tfrac{1}{2}(x_i - \theta_i 1_J)'\Gamma(x_i - \theta_i 1_J) + \beta_T T_i + \tfrac{1}{2}\beta_{TT} T_i^2 \\ &\quad + T_i (x_i - \theta_i 1_J)'\varphi + v_i \\ &\equiv \beta_0 + x_i'\beta + \tfrac{1}{2} x_i'\Gamma x_i + \beta_T T_i + \tfrac{1}{2}\beta_{TT} T_i^2 + x_i'\varphi T_i - g(\theta_i, x_i, T_i) + v_i, \end{aligned} \tag{3}$$

where y_i is the log of output, 1_J denotes the $J \times 1$ vector of ones, x_i is the $J \times 1$ vector of inputs in log, T_i is the trend/shift variable, v_i is statistical noise, β_0, β_T and β_{TT} are parameters, β, φ are $J \times 1$ parameter vectors, and Γ is a $J \times J$ symmetric matrix containing parameters. Finally, $g(\theta_i, x_i, T_i) = -[\tfrac{1}{2}\theta_i^2\Psi - \theta_i \Xi_i]$, where $\Psi = 1_J'\Gamma 1_J$, and $\Xi_i = 1_J'(\beta + \Gamma x_i + \varphi T_i)$, $i = 1,\ldots,n$. Note that Ψ is a constant while Ξ is a function of the data and parameters. The index i refers to a firm in a cross-sectional set up and can be used as an index for both firm and time in the case of panel data. Note that we re-parameterized Θ as $\Theta = \exp(-\theta)$ where θ is nonnegative. One can interpret θ (the IO technical inefficiency) as the percentage (when multiplied by 100) by which input-use is increased, for a given level of output, due to technical inefficiency.

The OO model, on the other hand, is

$$y_i = \beta_0 + x_i'\beta + \tfrac{1}{2} x_i'\Gamma x_i + \beta_T T_i + \tfrac{1}{2}\beta_{TT} T_i^2 + x_i'\varphi T_i - \lambda + v_i \qquad (4)$$

where Λ is re-parameterized as $\Lambda = \exp(-\lambda)$. The model in this form is the one introduced by Aigner et al. (1977) and since then been used by many, e.g., see Chapter 3 of Kumbhakar and Lovell (2000).

A close look at the specifications of the IO and OO models in (3) and (4), shows that the IO model can be converted to an OO model. Note that if the production function is homogeneous of degree *r*, then $\Gamma 1_J = 0, 1'_J \beta = r,$ and $1'_J \varphi = 0$. In such a case the $g(\theta_i, x_i)$ function becomes a constant multiple of θ, (viz., $[\frac{1}{2}\theta_i^2 \Psi - \theta \Xi_i]$ = $-r\theta_i$), and consequently, the IO model is indistinguishable from the OO model. Thus if the underlying production function is homogeneous, there is no need for new tools to estimate the IO production model. However, the IO model is econometrically different from the OO model when the production function is non-homogeneous. First, the component containing the inefficiency parameter, $g(\theta, x, T)$, depends on the data and the parameters. Thus the mean and variance of $g(\theta, x, T)$ will depend on the data and parameters. Furthermore, the parameters of the $g(\theta, x, T)$ function, other than those from the distribution of θ, are not new. They all come from the production function. Second, λ is often assumed to be distributed as a half-normal or a truncated normal variable with constant mean and variance. The same distributional assumptions can be made on θ but not on $g(\theta, x, T)$. Thus, econometric estimation of these two models will be different. Since in real world applications non-homogeneous production functions are quite common, it is necessary to develop tools for estimating IO technical inefficiency econometrically. We now focus our attention to the estimation of a non-homogeneous IO model that is never estimated in a primal framework.

Suppose that we have panel data and one is willing to make the assumption that IO technical inefficiency is time-invariant. We use this assumption as our starting point and we will relax it later. Then the model in (4) can be rewritten as

$$y_{it} = z_{it}'\alpha + \tfrac{1}{2}\theta_i^2 \Psi - \theta_i \Xi_{it} + v_{it}, \; i = 1,\ldots,n, \; t = 1,\ldots,T, \qquad (5)$$

where, as before, $\Psi = 1'_J \Gamma 1_J$, $\Xi_{it} = \delta + 1'_J(\beta + \Gamma x_{it} + T_{it}\varphi)$, $i = 1,...,n$, $t = 1,...,T$, θ_i is input-oriented technical inefficiency, and v_{it} is the (two-sided) error term. The *z* vector includes all the regressors (the *x*'s and *T*, squares and interactions of the *x*'s and *T*, and the intercept). This specification nests the case of cross-sectional data if we specify $T = 1$ so that the firm at different dates is considered as a different firm. So the model can be used for the case of panel as well as cross-sectional data. Suppose $v_{it} \sim IN(0, \sigma^2)$ and θ_i follows a distribution with the density function $p_\theta(\theta_i \mid \omega)$ where ω is an unknown parameter. We will specify this density function as exponential with mean $1/\omega$, i.e., $p_\theta(\theta_i \mid \omega) = \omega \exp(-\omega\theta_i)$ or alternatively as half-normal with scale parameter ω^2, viz., $p_\theta(\theta_i \mid \omega) = \left(\frac{\pi}{2}\omega^2\right)^{-1/2} \exp(-\theta_i^2 / 2\omega^2)$, and $\theta_i \geq 0$ in both cases.

Define $w_{it} = \frac{1}{2}\theta_i^2 \Psi - \theta_i \Xi_{it}$, which depends on the data, the parameters, and the θ_i's. Clearly, we have $y_{it} = z'_{it}\alpha + v_{it} + w_{it}$. Moreover, $E(y_{it} \mid z_{it}) = z'_{it}\alpha + E(w_{it} \mid z_{it})$, where $E(w_{it} \mid z_{it}) = \frac{1}{2}E(\theta_i^2)\Psi - E(\theta_i)\Xi_{it}$. Suppose $E(\theta_i) = \mu_\theta$ and $Var(\theta_i) = \sigma_\theta^2$, where both are functions of the parameter ω. It is easy to show that $E(w_{it} \mid z_{it}) = \frac{1}{2}(\mu_\theta + \sigma_\theta^2)\Psi - \mu_\theta \Xi_{it}$, and $E(y_{it} \mid z_{it}) = z'_{it}\alpha + \frac{1}{2}(\mu_\theta + \sigma_\theta^2)\Psi - \mu_\theta \Xi_{it}$.

Therefore, the least squares (LS) estimator of all the parameters (not just the intercept) will be inconsistent because the frontier is not shifted neutrally. In other words, the regression function given by $E(y_{it} \mid z_{it})$ is not simply $z'_{it}\alpha$ up to a constant. The implication is that we cannot apply corrected LS or similar techniques, as done in the familiar output-oriented frontier models. Consequently, measures of input elasticities, returns to scale and technical change based on the LS estimator will also be inconsistent. To get proper estimates of these, one has to use the ML method.

The likelihood function of the model is given by

$$L(\alpha,\sigma,\omega,\theta; y,Z) = (2\pi\sigma^2)^{-nT/2} \prod_{i=1}^{n} \int_0^\infty \exp[-\tfrac{1}{2}\sum_{t=1}^{T} (y_{it} - z'_{it}\alpha - \tfrac{1}{2}\theta_i^2\Psi + \theta_i\Xi_{it})^2 / \sigma^2] p_\theta(\theta_i \mid \omega) d\theta_i \quad (6)$$

The integral with respect to the latent θ_i is not available in closed form. The implication of this is that both the sampling-theory and the Bayesian approach have to rely on using numerical techniques. In this paper, we follow a Bayesian approach. According to Bayes' theorem the kernel posterior density function is given by $p(\alpha,\sigma,\omega \mid y,Z) \propto L(\alpha,\sigma,\omega;y,Z)p(\alpha,\sigma,\omega)$ where $p(\alpha,\sigma,\omega)$ is the prior. We use the data augmentation scheme. In particular, we consider the θ_i's as parameters. Consequently, use of Bayes' theorem gives the following augmented kernel posterior density function $p(\alpha,\sigma,\omega,\theta \mid y,Z) \propto L(\alpha,\sigma,\omega,\theta;y,Z)p(\alpha,\sigma,\omega,\theta)$, where $\theta = [\theta_1,...,\theta_n]'$. Specifically, we have

$$L(\alpha,\sigma,\omega,\theta;y,Z) = (2\pi\sigma^2)^{-nT/2}\prod_{i=1}^{n} \exp[-\tfrac{1}{2}\sum_{t=1}^{T} (y_{it} - z_{it}'\alpha - \tfrac{1}{2}\theta_i^2\Psi + \theta_i\Xi_{it})^2/\sigma^2]p_\theta(\theta_i \mid \omega) \quad (7)$$

and

$$p(\alpha,\sigma,\omega,\theta) = p(\alpha,\sigma,\omega)\prod_{i=1}^{n} p_\theta(\theta_i \mid \omega). \quad (8)$$

3.2. Priors

We adopt a prior distribution of the form

$$\alpha \sim N\left(\alpha_0,\ V_0^{-1}\right)$$

where α_0 is the prior mean and V_0 is the prior precision matrix. To specify the prior precision we choose Zellner's g-prior with $V_0^{-1} = g(Z'Z)^{-1}$, where α_0 is a vector of zeros. The advantage of the g-prior is that it requires specification of a single, scalar parameter. Specifying the prior precision matrix in normal priors necessitates knowledge about the covariances of parameters which are, usually, not available. Other than that the techniques proposed in this paper can be easily modified to accommodate arbitrary multivariate normal priors. We choose $g = 10^4$ to have an informative but very diffuse prior. We also assume the following priors for parameters σ^2 and ω^2:

$$\frac{\overline{Q}}{\sigma^2} \mid \cdot \sim \chi^2(\overline{n}),$$

$\dfrac{Q_0}{\omega^2} \mid \cdot \sim \chi^2(n_0)$ in the half-normal model, and

$\omega \sim Gamma(n_0, Q_0)$ in the exponential model,

where $\overline{n}$, n_0, $\overline{Q}$, $Q_0 \geq 0$ are parameters of the prior distributions. We specify $\overline{n} = 1$, and $\overline{Q} = 10^{-4}$, which is very diffuse but proper. It remains to specify the values of prior parameters n_0 and Q_0. To make the priors as diffuse as possible we set $n_0 = 1$, and we choose the value of Q_0 to match prior *median* efficiency with pre-assigned values, viz., 0.90, 0.75 and 0.50. We consider these values to be reasonable when sensitivity analysis with respect to prior assumptions about efficiency is an issue. Therefore, we have three priors that we call I, II and III. To determine the value of Q_0 we need to solve a nonlinear equation where the prior median efficiency for each model is computed with simulation. The procedure for the exponential model is as follows. We solve the nonlinear equation $M(Q_0) = \overline{M}$, where $\overline{M}$ is one of the pre-assigned values for prior median efficiency, and the function $M(Q_0)$ is defined as follows.

1. We simulate N random numbers $\omega_{(s)} \sim Gamma(n_0, Q_0)$, $s = 1,...,S$, $n_0 = 1$.
2. We simulate inefficiency draws $u_{(s)}$ from an exponential distribution with mean $1/\omega_{(s)}$, and we obtain the efficiency draws, $r_{(s)} = \exp(-u_{(s)})$, $s = 1,...,S$.
3. We compute the median of $\{r_{(s)}, s = 1,...,S\}$, which is $M(Q_0)$.
4. Return the value $M(Q_0)$.

To solve the nonlinear equation $M(Q_0) = \overline{M}$, we apply Newton's method with numerical derivatives. We find that convergence is rapid for the given $\overline{M}$, when $S = 100$ and the results are not particularly sensitive to higher values of S. The prior parameters derived from the solution of the nonlinear equation, are presented in Table 1. The table is used as follows. For prior specification I, median efficiency is 0.90. In that case the parameter Q_0 should be 0.0165 if we choose the half-normal model, and 0.0307 if we choose the exponential model. For different values of prior median efficiency, we choose prior II or III and we obtain the corresponding values of Q_0.

Table. Prior parameter values.

Prior and median efficiency	Q_0, half-normal	Q_0, exponential
I, 0.90	0.0165	0.0307
II, 0.75	0.123	0.149
III, 0.50	0.715	0.661

3.3. Bayesian computation

Based on the augmented kernel posterior distribution we can use Gibbs sampling with data augmentation to perform the Bayesian computations, and provide marginal posterior moments and densities (For details see Gelfand and Smith (1990) and Tanner and Wong (1987)). To implement Gibbs sampling we need to draw random numbers from the posterior conditional density functions, and we have to do this in a computationally efficient way. We turn attention to this matter in what follows.

The conditional posterior density function of θ_i is given by

$$p(\theta_i \mid \cdot) \propto \exp\left[-\frac{\sum_{t=1}^{T} \left(\tfrac{1}{2}\theta_i^2 \Psi - \Xi_{it}\theta_i - R_{it} \right)^2}{2\sigma^2} - \frac{\theta_i^2}{2\omega^2} \right], \ \theta_i \geq 0, \ i = 1,\ldots,n, \qquad (9)$$

where $R_{it} = y_{it} - z_{it}'\alpha$. To generate random draws from the distribution whose kernel density function is given by (9), suppose $\Psi \approx 0$. In that case the conditional posterior of θ_i is approximately truncated normal, *i.e.*,

$$\theta_i \mid \cdot \sim N_+\left(m_i,\ V_i^2\right) \text{ approximately,}$$

where $m_i = -\dfrac{\omega^2 \Xi_i' R_i}{\omega^2 \Xi_i' \Xi_i + \sigma^2}$, $V_i^2 = \dfrac{\sigma^2 \omega^2}{\omega^2 \Xi_i' \Xi_i + \sigma^2}$, $\Xi_i = [\Xi_{i1},\ldots,\Xi_{iT}]'$, $R_i = [R_{i1},\ldots,R_{iT}]'$ are $T \times 1$ vectors. Denote by $g(\theta_i) = (2\pi V_i^2)^{-1/2} \exp\left[-\dfrac{(\theta_i - m_i)^2}{2V_i^2} \right] \Phi(m_i / V_i)^{-1}$ the density of a random variable with the $N_+\left(m_i,\ V_i^2\right)$ distribution. To generate a random draw from the conditional distribution we generate a candidate draw $\tilde{\theta}_i \sim N_+\left(m_i, V_i^2\right)$, and we accept with probability $\min\left(1, \frac{p(\tilde{\theta}_i \mid \cdot)/g(\tilde{\theta}_i)}{p(\theta_i \mid \cdot)/g(\theta_i)}\right)$ where θ_i is the current draw. To generate a candidate draw from the truncated normal distribution we use the acceptance procedure described in Tsionas (2000).

To draw from the conditional distribution of parameters α we use the representation in (3). Define $x_{it}^* = x_{it} - \theta_i$ conditional on the value of θ_i. Model (3) may be written in the form

$$y = Z\alpha + v$$

where $z_{it} = \begin{bmatrix} x_{it}^* \\ vech(x_{it}^* \otimes x_{it}^*) \\ T_{it} x_{it}^* \end{bmatrix}$.

The conditional posterior distribution of α is

$$\alpha \mid \cdot \sim N(\tilde{\alpha}, \tilde{V}),$$

where $\tilde{\alpha} = (Z'Z + \sigma^2 V_0)^{-1}(Z'y + \sigma^2 V_0 \alpha_0)$, and $\tilde{V} = \sigma^2 (Z'Z + \sigma^2 V_0)^{-1}$.

For the g-prior these expressions simplify to

$$\tilde{\alpha} = \left(\frac{g}{g+\sigma^2}\right) a + \left(\frac{\sigma^2}{g+\sigma^2}\right)\alpha_0,$$

where $a = (Z'Z)^{-1} Z'y$ is the LS estimator, and

$$\tilde{V} = \frac{g\sigma^2}{g+\sigma^2}(Z'Z)^{-1}.$$

The posterior conditional distributions of the scale parameters are

$$\frac{\overline{Q} + (y - Z\alpha)'(y - Z\alpha)}{\sigma^2} \mid \cdot \sim \chi^2(\overline{n} + nT),$$

$$\frac{Q_0 + \theta'\theta}{\omega^2} \mid \cdot \sim \chi^2(n_0 + n).$$

The alternative specification is an exponential distribution for the θ_i s, viz.,

$$p(\theta_i \mid \omega) = \omega \exp(-\omega\theta_i)\,,\ \omega > 0\,,\ \theta_i \geq 0\,. \tag{10}$$

In this case we have $1/\omega = E(\theta_i \mid \omega)$. Under this specification, the conditional posterior density function of θ_i is given by

$$p(\theta_i \mid \cdot) \propto \exp\left[-\frac{\sum_{t=1}^{T}\left(\frac{1}{2}\theta_i^2\Psi - \Xi_{it}\theta_i - R_{it}\right)^2}{2\sigma^2} - \omega\theta_i \right],\ \theta_i \geq 0\,, \tag{11}$$

and we can generate random draws by following similar principles. When $\Psi \approx 0$, we have

$\theta_i \mid \cdot \sim N_+(m_i,\ V_i)$ approximately,

where $m_i = -\dfrac{\Xi_i' R_i + \omega\sigma^2}{\Xi_i'\Xi_i}$, and $V_i = \dfrac{\sigma^2}{\Xi_i'\Xi_i}$,

and all other variables have been defined before. The conditional posterior distribution of ω is $Gamma(n + n_0,\ Q_0 + \theta' 1_n)$ assuming that the prior is $Gamma(n_0, Q_0)$, where $n_0, Q_0 \geq 0$. Drawing from the conditional posterior distributions of the other parameters is exactly the same as in the half-normal case so Bayesian analysis based on Gibbs sampling can be implemented efficiently.

The previous model can be applied to *cross-sectional data* if we assume $T = 1$ so that the firm at different periods is treated as a different firm. The likelihood function, the kernel posterior distribution, and the posterior conditional distributions can be easily derived by using $T = 1$ in the relevant expressions. More specifically, the model is

$$y_i = z_i'\alpha + \tfrac{1}{2}\theta_i^2\Psi - \theta_i\Xi_i + v_i\,,\ i = 1,\ldots,n$$

where $\Xi_i = \delta + 1'_J(\beta + \Gamma x_i)$ as in equation (5). Of course, with cross-sectional data it is not possible to have a trend variable in the model. The likelihood function is

$$L(\alpha,\sigma,\omega,\theta;y,Z) = (2\pi\sigma^2)^{-nT/2}\prod_{i=1}^{n}\int_0^\infty \exp[-\tfrac{1}{2}(y_i - z_i'\alpha - \tfrac{1}{2}\theta_i^2\Psi + \theta_i\Xi_i)^2/\sigma^2]p_\theta(\theta_i \mid \omega)d\theta_i .$$

Adopting the same set of priors like before, we use Gibbs sampling with data augmentation with the following conditional distributions. The conditional posterior distribution of θ_i under the half-normal specification is given by

$$p(\theta_i \mid \cdot) \propto \exp\left[-\frac{\left(\frac{1}{2}\theta_i^2\Psi - \Xi_i\theta_i - R_i\right)^2}{2\sigma^2} - \frac{\theta_i^2}{2\omega^2}\right],\ \theta_i \geq 0,\ i = 1,\ldots,n,$$

where $R_i = y_i - z_i'\alpha$. Under the exponential specification, the conditional distribution becomes

$$p(\theta_i \mid \cdot) \propto \exp\left[-\frac{\left(\frac{1}{2}\theta_i^2\Psi - \Xi_i\theta_i - R_i\right)^2}{2\sigma^2} - \omega\theta_i\right].$$

Draws from these distributions can be obtained by using straightforward extensions of the techniques developed for the case of panel data. The conditionals of α, σ, and ω are as follows. We can write the model as

$$y = Z\alpha + v,$$

where $z_i = \begin{bmatrix} x_i^* \\ vech(x_i^* \otimes x_i^*) \end{bmatrix}$.

The conditional posterior distribution of α is

$$\alpha \mid \cdot \sim N(\tilde{\alpha}, \tilde{V}),$$

where $\tilde{\alpha} = (Z'Z + \sigma^2 V_0)^{-1}(Z'y + \sigma^2 V_0 \alpha_0)$, $\tilde{V} = \sigma^2 (Z'Z + \sigma^2 V_0)^{-1}$.

For the g-prior these expressions simplify to

$$\tilde{\alpha} = \left(\frac{g}{g+\sigma^2} \right) a + \left(\frac{\sigma^2}{g+\sigma^2} \right) \alpha_0 ,$$

where $a = (Z'Z)^{-1} Z'y$ is the LS estimator, and $\tilde{V} = \frac{g\sigma^2}{g+\sigma^2}(Z'Z)^{-1}$.

The posterior conditional distributions of the scale parameters are

$$\frac{\overline{Q} + (y - Z\alpha)'(y - Z\alpha)}{\sigma^2} | \cdot \sim \chi^2(\overline{n} + n) ,$$

$$\frac{Q_0 + \theta'\theta}{\omega^2} | \cdot \sim \chi^2(n_0 + n) .$$

As in the case of panel data, random number generation from these distributions is straightforward.

4. Efficiency measurement

Measures of efficiency can be obtained using the procedure developed by van den Broeck, Koop, Osiewalski, and Steel (1994) with the necessary modifications to accommodate our input-oriented framework. First, we define *posterior predictive input-oriented technical efficiency*. Under the half-normal model $\theta_i | \omega \sim N_+(0,\ \omega^2)$ so we can easily determine the distribution of $r_i = \exp(-\theta_i)$ given ω, and we can integrate out parameter uncertainty to find the distribution with density $p(r_i | y, X) = \int p(r_i | \omega) p(\omega | y, X) d\omega$, where $p(\omega | y, X)$ is the marginal posterior density function of ω. This is the distribution of input-oriented technical efficiency for a "typical" or yet-unobserved firm. Similar principles can be adopted to define the posterior predictive input-oriented technical efficiency under the exponential specification. *Firm - specific input-oriented technical efficiency* can be defined as $F_i = S^{-1} \sum_{s=1}^{S} \theta_i^{(s)}$, where $\theta_i^{(s)}$ is the s th Gibbs draw from the conditional distribution of $\theta_i^{(s)}$ given the other parameters and the data.

5. Empirical results

We used data on 80 Spanish dairy farms observed from 1993 to 1998. These are small family farms operated mostly by the family members. We use liters of milk as output and number of cows, kilograms of concentrates, hectares of land and labor (measured in man-equivalent units) as inputs (see Alvarez, Arias, and Kumbhakar (2003) for details). Since there are only 80 farms in the data in our cross-sectional model we use all observations without using any panel feature of the data and treat the data as a single cross-section. In doing so we treat each observation as a separate farm. However, we used time as an additional regressor to capture technical change. Thus, our IO model is the one specified in equation (5) in which the *x* variables are the four inputs, and *T* is the time trend variable. The OO model is given in (6). In the panel data model we use the same specification, except that technical inefficiency (both in the IO and OO models) is assumed to be time-invariant. To make this assumption less onerous we used the data for only 3 years (1993-1995). Thus, a total of 240 observations are used for both the cross-sectional and panel models.

5.1 Cross-sectional results

Before examining characteristics of the technology such as the returns to scale (RTS) under different models, we note that RTS (defined as RTS = $\sum_j \partial y / \partial x_j$) is not affected by the presence of technical inefficiency in the OO model. The same is true for input elasticities and elasticities of substitution (that are not explored here). This is because inefficiency in the OO model shifts the production function in a neutral fashion. On the contrary, the magnitude of technical inefficiency affects input elasticities and RTS in the IO models. Using the translog specification in (5) we get

$$RTS\_IO = 1'_J(\beta + \Gamma x_i + \varphi T_i) - \theta_i 1'_J \Gamma \qquad (12)$$

whereas the formula for RTS in the OO model is

$$RTS\_OO = 1'_J(\beta + \Gamma x_i + \varphi T_i) \qquad (13)$$

It is often argued that RTS and other characteristics of the technology should be defined at the frontier. If so the formula for calculating RTS in the IO and OO models will be the same.

Similarly, if the production function is homogeneous (i.e., $1_J^{'}\Gamma = 0$), the two definitions will be identical. Empirically, some differences are likely to be observed because the estimated parameters may not be exactly the same in the IO and OO models.

Similar to RTS, TC in the IO model can be measured conditional on θ (TC_IO(I)) and TC defined at the frontier (TC_IO(II)), viz.,

$$\text{TC\_IO(I)} = \beta_T + \beta_{TT} T_i + x_i^{\prime}\varphi + 1_J^{'}\varphi\theta_i \tag{14}$$

$$\text{TC\_IO(II)} = \beta_T + \beta_{TT} T_i + x_i^{\prime}\varphi \tag{15}$$

These two formulae will give different results unless technical change is either neutral or the production function is homogeneous (i.e., $1_J^{'}\varphi \neq 0$). The formula for TC_OO is the same as TC_IO(II), except that the estimated parameters in (15) are from the IO model whereas the parameters to compute TC_OO are from the OO model. The estimated parameters from the two models are likely to differ as well.

In our application we failed to accept the hypothesis that the production function is homogeneous, using the standard likelihood ratio test. Thus, it is likely that there will be some observable differences in the estimated RTS. We consider two measures of RTS in the IO model (type I RTS uses the formula in (12) while type II RTS is defined at the frontier) and compare them with the RTS_OO from (13). Note that RTS_OO and type II RTS are based on exactly the same formula but the parameters in (12) are obtained from estimating the IO model in (5) while those in (13) are obtained from estimating the OO model in (6). Thus, their differences, if any, will be due to differences in parameter estimates.

Returns to scale and technical change measures (type I and type II) for the three priors are reported in Figures 1 and 2. Figures 1a and 1b provide results for returns to scale, and Figures 1c and 1d for technical change for type I. Type II measures are reported in Figure 2 for the three different priors. The measures are firm-specific and parameter uncertainty is averaged out in standard Bayesian fashion, *i.e.*, the measures are averaged against posterior draws from the MCMC simulation. These results are reasonably robust with respect to the different priors although technical change measures are clearly less sensitive. Returns to scale are, for the most part, increasing and, on the average, close to 1.2. Moreover, the half-normal and exponential models give comparable results.

We noted before that TE in the IO and OO models (i.e., Λ and Θ) are not the same unless the returns to scale are unity (in which case the percentage decrease in inputs and output are same). To make these measures comparable, we convert the IO measure in terms of the OO measure using the formula TE_OO(CO) = exp($-g(\theta, x, T)$), where $g(\theta, x, T)$ is defined beneath equation (4). Since both measures are ratios of actual to frontier output we can compare them directly. The distributions of these two efficiency measures (each for half normal and exponential distributions) are plotted in Figure 3. The upper panel provides the posterior predictive input-oriented efficiency distributions for the half-normal (left) and exponential (right) models for all three priors. These distributions are somewhat sensitive to the priors since we treat the panel as a cross-section in this analysis. In all cases, however, the mass of the posterior predictive distributions is above 0.80 suggesting that values lower than 80% are less likely for the data. Kernel densities of firm-specific input-oriented efficiency measures are reported in the middle panel. Again, these distributions are somewhat sensitive to the nature of the priors. For the half-normal model, the means are close to 92%. For the exponential model the means are somewhat higher and close to 95%. Other than this difference in location, the half-normal and exponential models are close in terms of the shape of distributions for firm-specific input-oriented efficiency. In the lower panel of Figure 1, we reported the kernel densities of implied output-oriented efficiency for the two models and three priors. The results are somewhat sensitive to the prior, the general shape of distributions is about the same in the two models, and output-based efficiency is somewhat higher in the exponential model. Average output-based firm-specific efficiency is above 90% in both models.

In Figure 4 we provide returns to scale and technical change measures based on a Bayesian analysis of the ALS model using the same priors as before and both the half-normal and exponential distributional assumptions on technical inefficiency. The results are somewhat different compared to Figures 1 and 2, especially in that the ALS results do not display the strong bimodality of distributions obtained from the IO models.

Posterior predictive and firm-specific output-oriented efficiency distributions derived from the ALS models are reported in Figure 5. Average firm-specific efficiency is close to 90% for both the half-normal and exponential models, the results are very similar in the two models, and less sensitive to priors compared with the IO results.

In Figure 6, we provide much of the same information in a different way to facilitate cross-model comparisons. In the upper panel we present distributions of returns to scale, technical change, and output-based efficiency associated with the first prior for the IO (types I and II) and the ALS models (based on half-normal distribution). In the lower panel we report corresponding results for the model based on an exponential distribution. Returns to scale and technical change (type I and type II) measures are very close but differ somewhat from the ALS model. The differences are quite distinct when it comes to comparing distributions of output-based efficiency from the IO and OO models (Figures 6c and 6f). Contrary to the IO models, the OO models have a much larger left tail suggesting that lower efficiency is quite likely. For example, in the half-normal model, the IO model implies that efficiency levels less than 80% are highly unlikely, whereas according to the OO model firm-specific efficiency can be as low as 64%. These figures are about 90% and 74% in the exponential specification.

5.2 Results from panel data

Next, we consider an analysis based on panel data. We assume time-invariant input-oriented technical inefficiency and consider two alternative values of T (viz., $T = 2$ and $T = 3$). We report the results in Figure 7 in the form of kernel densities of firm-specific input-oriented technical inefficiency measures. The upper panel for all three priors is based on the half-normal specification. The lower panel reports results for the exponential specification. The results are sensitive to whether we assume cross-sectional or panel data but less sensitive to the value of T. Corresponding output-based efficiency distributions are reported in Figure 8. The most important differences between cross-sectional and panel data specifications, is that cross-sectional models imply much larger efficiency. With a panel data approach, it appears much more likely to have lower efficiency, and the panel-based distributions do not have the characteristic peak of the cross-sectional efficiency distributions at efficiency values close to either 0.94 (half-normal) or 0.98 (exponential).

From these results it is clear that the IO and OO models are different when it comes to comparing returns to scale or technical change and more importantly they have quite different implications when it comes to output-based efficiency. It also makes a difference whether we have cross-sectional or panel data. The differences implied by half-normal and exponential models are small, and the same is true for the efficiency priors we have considered.

5. Conclusions

This paper developed a Bayesian method for estimating input-oriented (IO) technical inefficiency using stochastic frontier production models. A flexible production function is used for this. Although the IO model is always used in the context of a stochastic cost function, it is never estimated from stochastic production frontiers in a cross-sectional setting. We used both cross-sectional and panel model specifications to a sample of Spanish dairy farms. Results from these models are compared to those from the standard output-oriented (OO) stochastic frontier model with both normal-half normal and normal-exponential cases.

We emphasize that our focus is not on the choice of efficiency orientation but on estimation and the empirical consequences of using different efficiency orientations. Our results suggest that the estimated technology (parameters), technical efficiency, returns to scale, technical change, *etc*., are sensitive to the choice of IO and OO models. This is true for both cross-sectional and panel models.

References

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.

Alvarez, A., C. Arias, and S.C. Kumbhakar, 2003, Empirical consequences of direction choice in technical efficiency analysis, manuscript, SUNY Binghamton, NY.

van den Broeck, J., G. Koop, J. Osiewalski, and M.F.J. Steel, 1994, Stochastic frontier models: A Bayesian perspective, Journal of Econometrics 61, 273-303.

Fare, R. and C.A.K. Lovell, 1978, Measuring technical efficiency in production, Journal of Economic Theory, 19, 150-162.

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.

Koop, G, J. Osiewalski, and M.F.J. Steel, 1997, Bayesian efficiency analysis through individual effects: Hospital cost frontiers, Journal of Econometrics 76: 77-105.

Koop, G., M.F.J. Steel, and J. Osiewalski, 1995, Posterior analysis of stochastic frontier models using Gibbs sampling, Computational Statistics 10, 353-373.

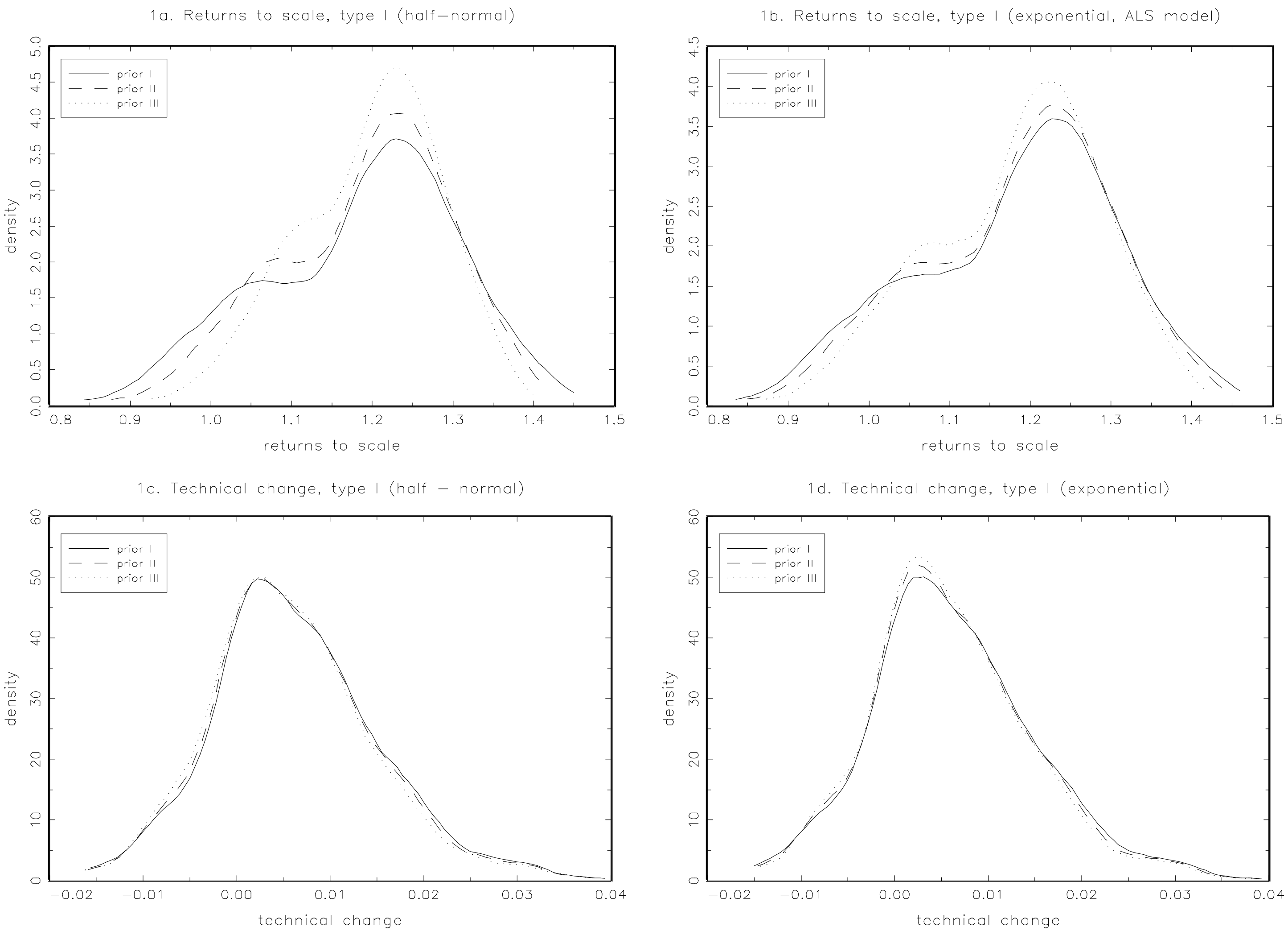
Kumbhakar, S.C. and C.A.K Lovell, 2000, Stochastic frontier analysis (Cambridge University Press, New York).

Meeusen, W., J. van den Broeck, 1977, Efficiency estimation from Cobb-Douglas production functions with composed error, International Economic Review 8, 435-444.

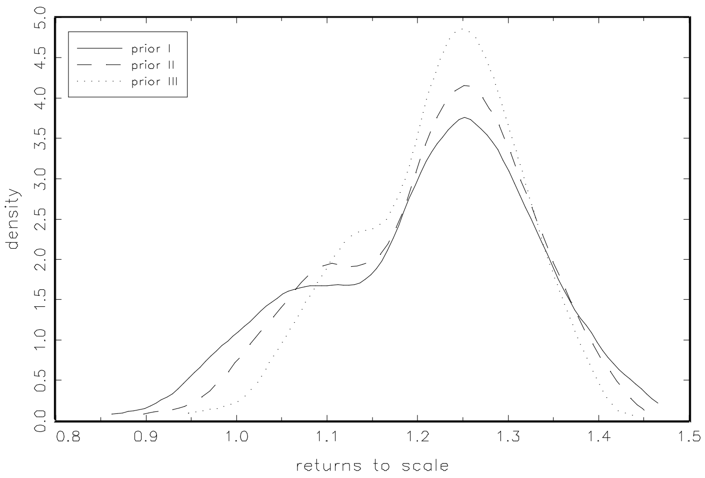
Ray, S.C. (2003), *Data Envelopment Analysis*, Cambridge University Press, New York (*forthcoming*).

Tanner, M.A. and W.H. Wong, 1987, The calculation of posterior distributions by data augmentation. Journal of the American Statistical Association 82, 528-550.

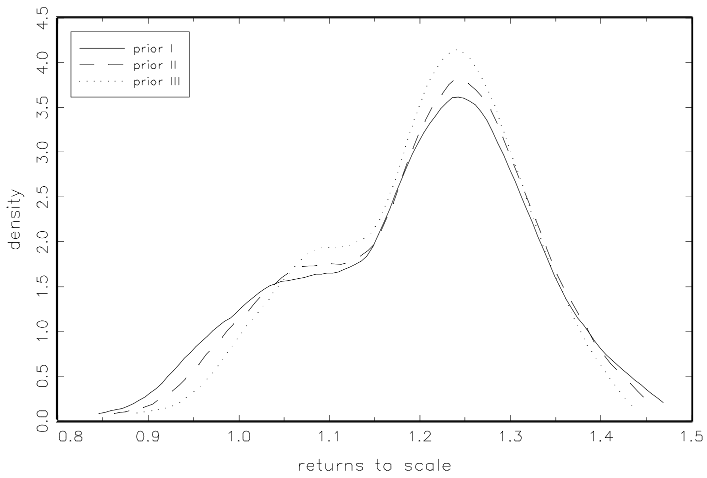
Tsionas, E.G., 2000, Full likelihood inference in normal-gamma stochastic frontier models, Journal of Productivity Analysis 13, 179-201.


1a. Returns to scale, type I (half−normal)
density
returns to scale
prior I
prior II
prior III
1b. Returns to scale, type I (exponential, ALS model)
density
returns to scale
prior I
prior II
prior III
1c. Technical change, type I (half − normal)
density
technical change
prior I
prior II
prior III
1d. Technical change, type I (exponential)
density
technical change
prior I
prior II
prior III

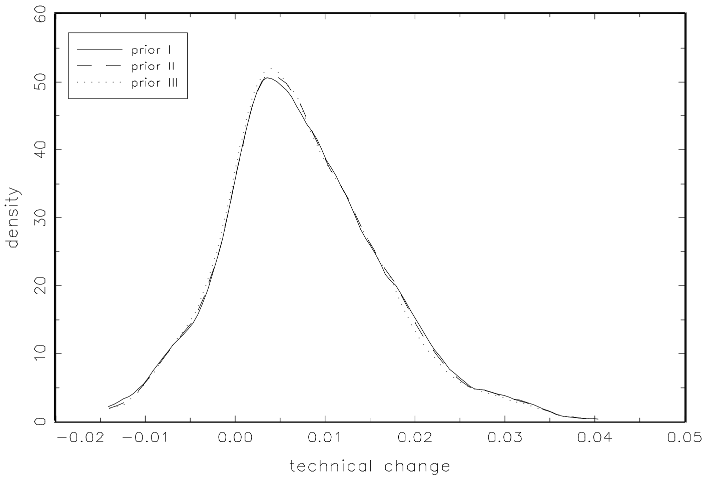
2a. Returns to scale, type II (half−normal)



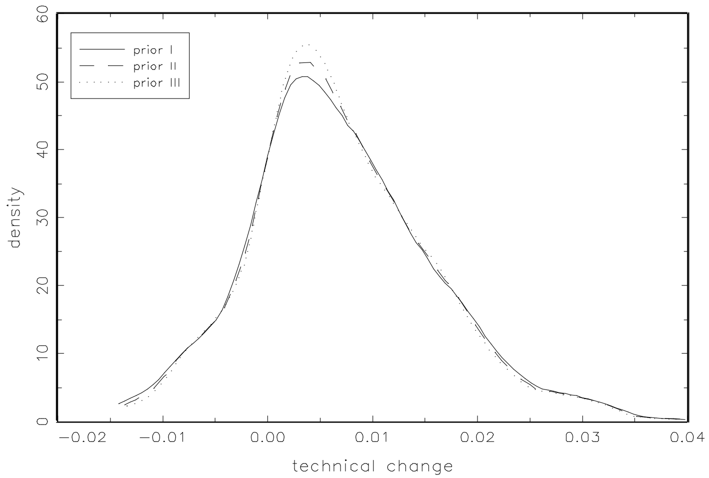
2b. Returns to scale, type II (exponential)

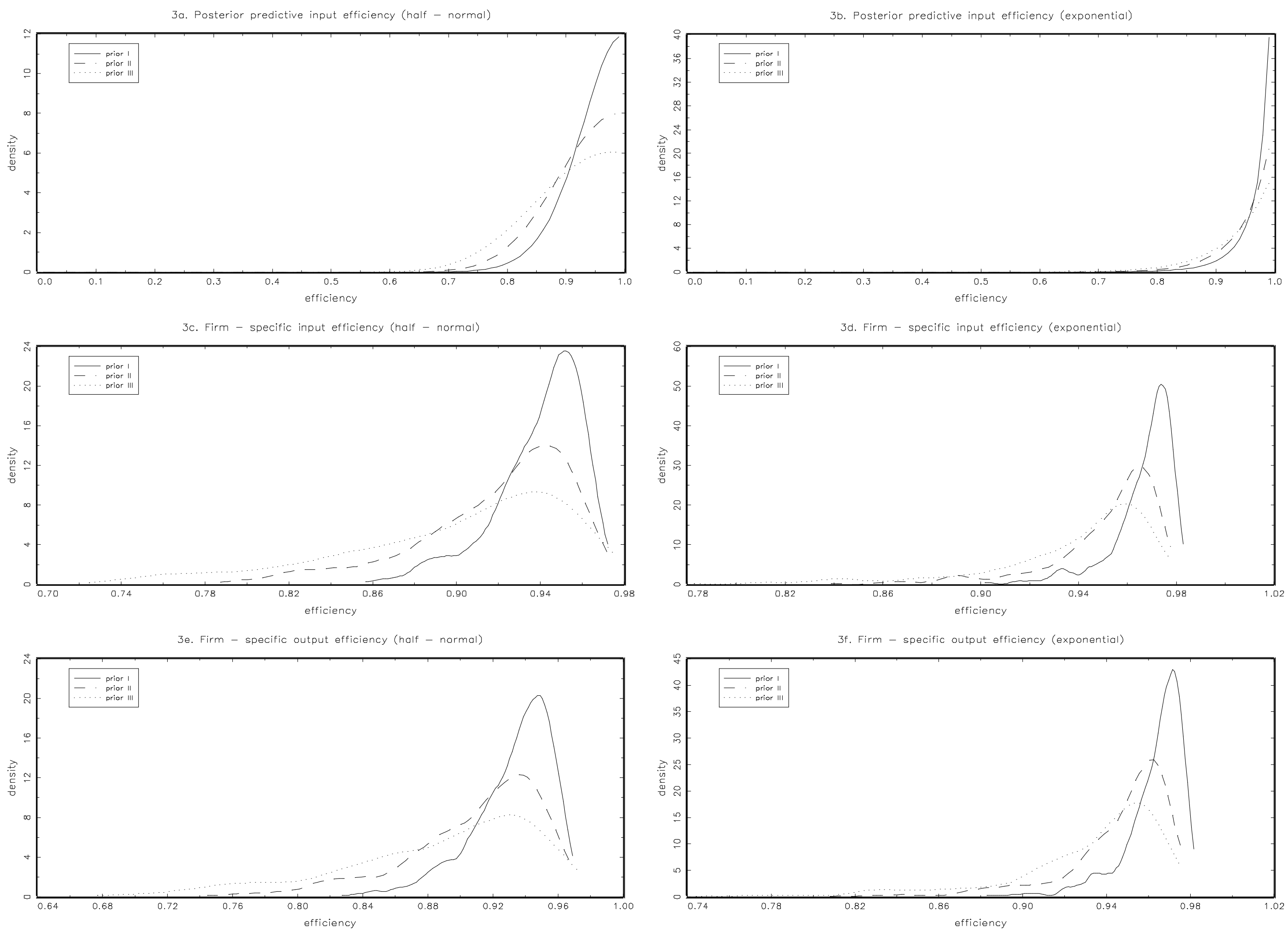


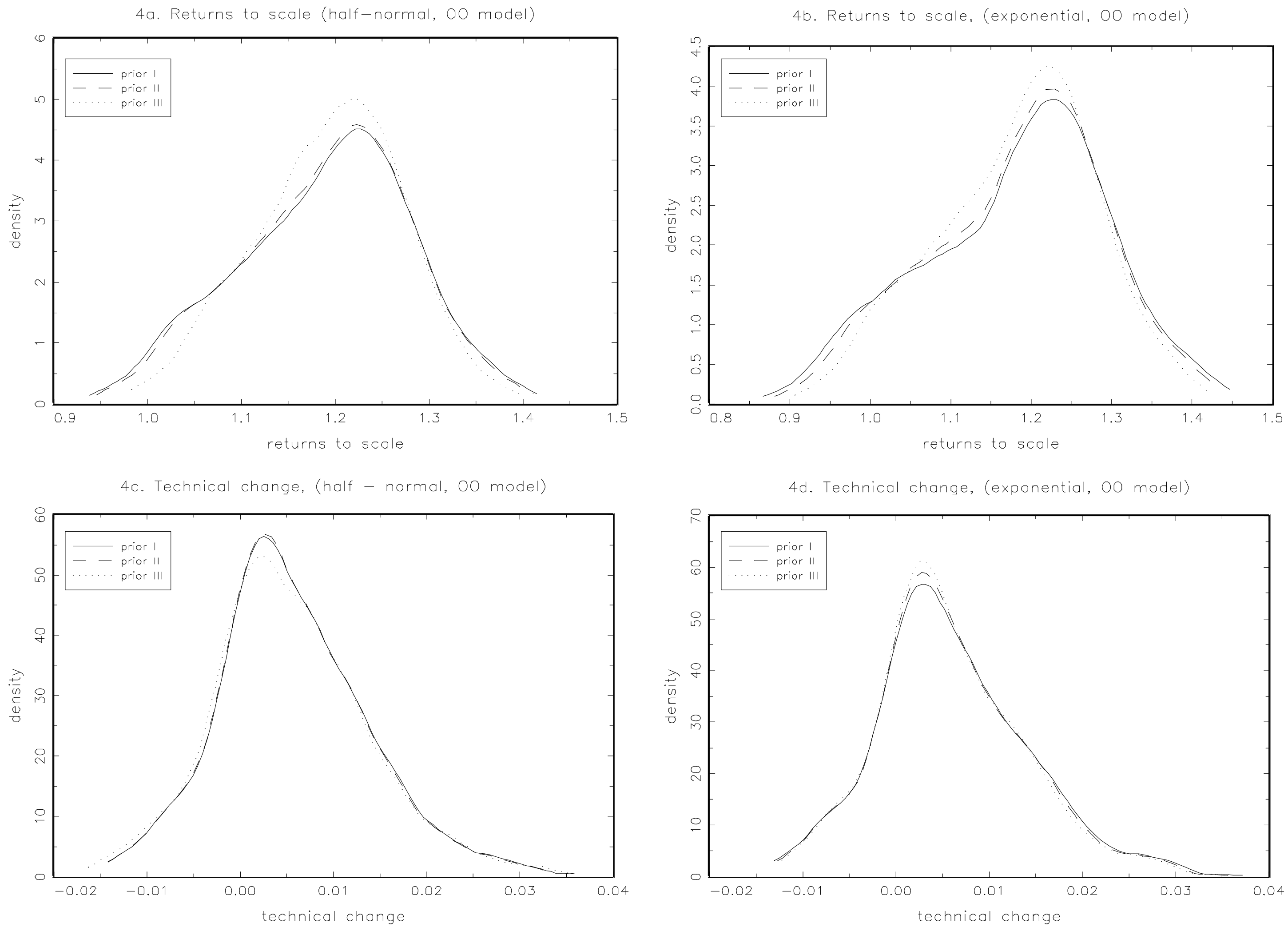
2c. Technical change, type II (half − normal)



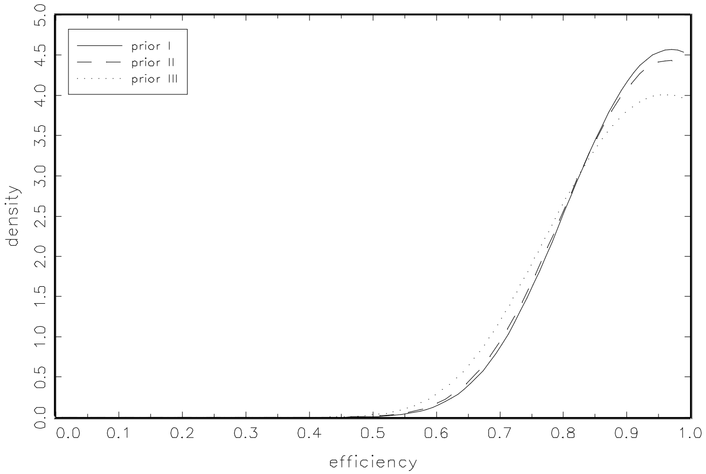
2d. Technical change, type II (exponential)



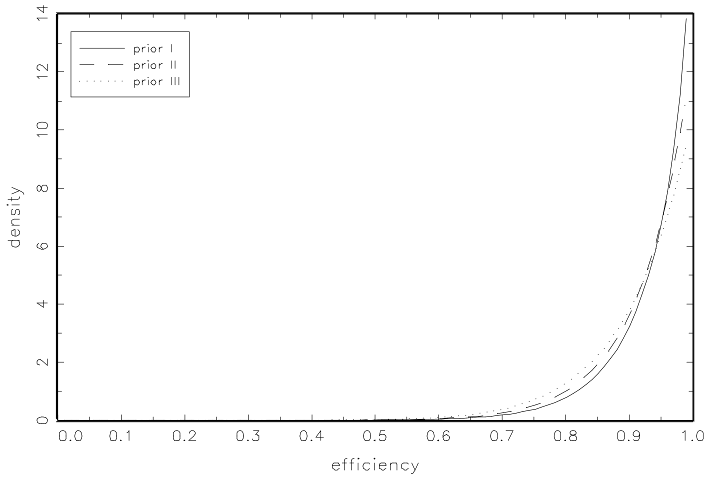

3a. Posterior predictive input efficiency (half – normal)
3b. Posterior predictive input efficiency (exponential)
3c. Firm – specific input efficiency (half – normal)
3d. Firm – specific input efficiency (exponential)
3e. Firm – specific output efficiency (half – normal)
3f. Firm – specific output efficiency (exponential)
prior I
prior II
prior III
density
efficiency


4a. Returns to scale (half-normal, OO model)
density
returns to scale
prior I
prior II
prior III
0
1
2
3
4
5
6
0.9
1.0
1.1
1.2
1.3
1.4
1.5
4b. Returns to scale, (exponential, OO model)
density
returns to scale
prior I
prior II
prior III
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
0.8
0.9
1.0
1.1
1.2
1.3
1.4
1.5
4c. Technical change, (half – normal, OO model)
density
technical change
prior I
prior II
prior III
0
10
20
30
40
50
60
−0.02
−0.01
0.00
0.01
0.02
0.03
0.04
4d. Technical change, (exponential, OO model)
density
technical change
prior I
prior II
prior III
0
10
20
30
40
50
60
70
−0.02
−0.01
0.00
0.01
0.02
0.03
0.04

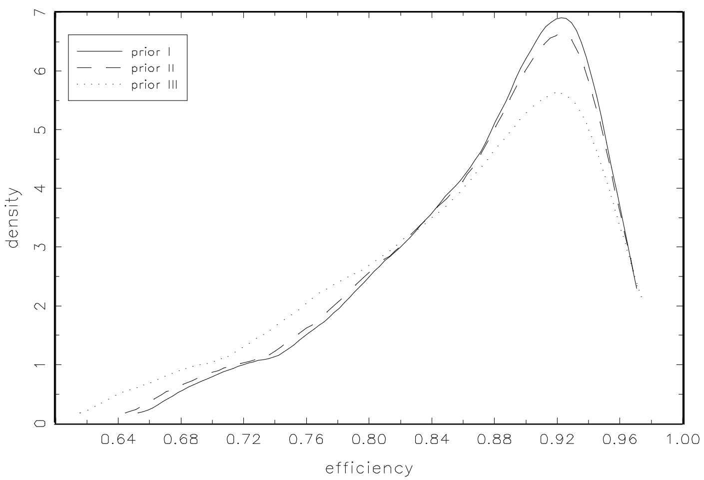
5a. Posterior predictive efficiency (half − normal, OO model)



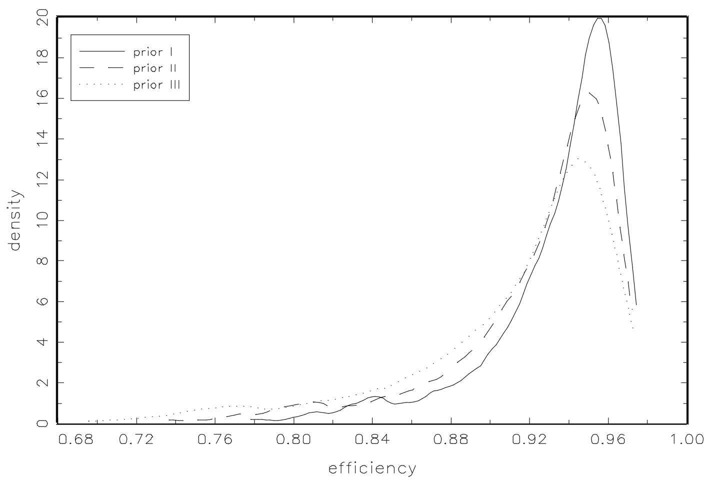
5b. Posterior predictive efficiency (exponential, OO model)



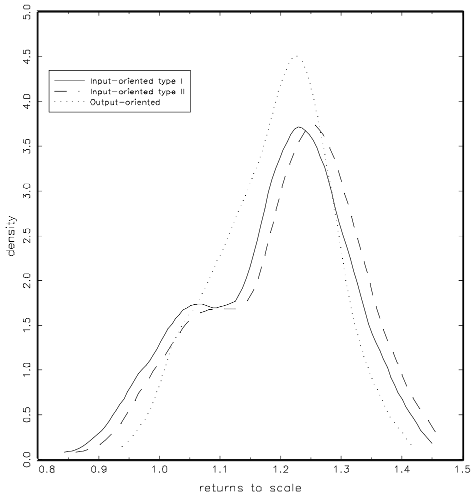
5c. Firm − specific efficiency (half − normal, OO model)



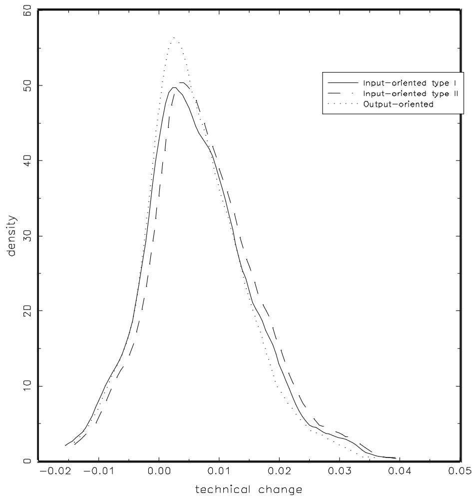
5d. Firm − specific efficiency (exponential, OO model)



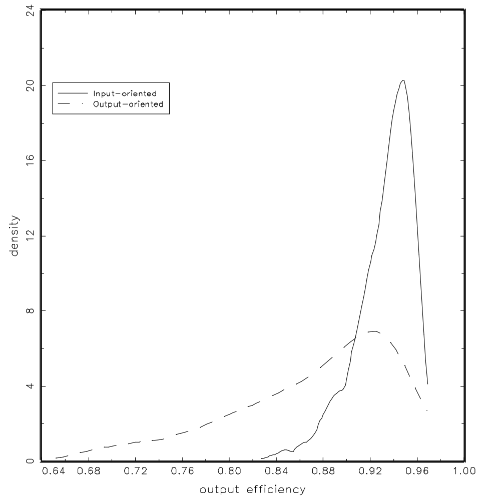
6a. Returns to scale (half-normal, prior I)



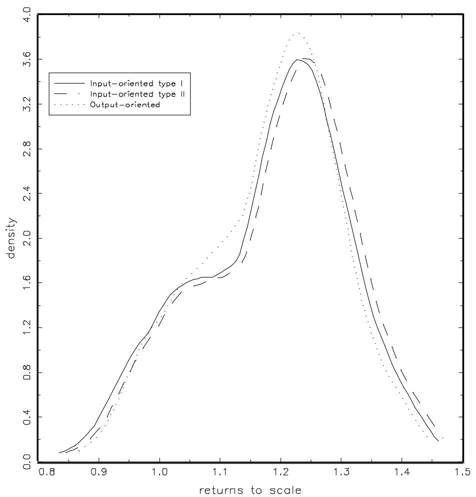
6b. Technical change (half-normal, prior I)



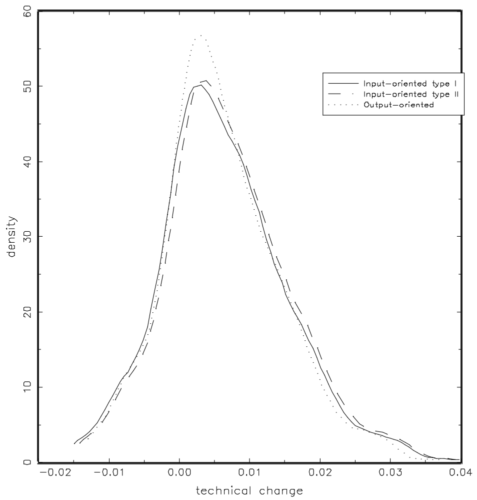
6c. Output efficiency (half-normal, prior I)



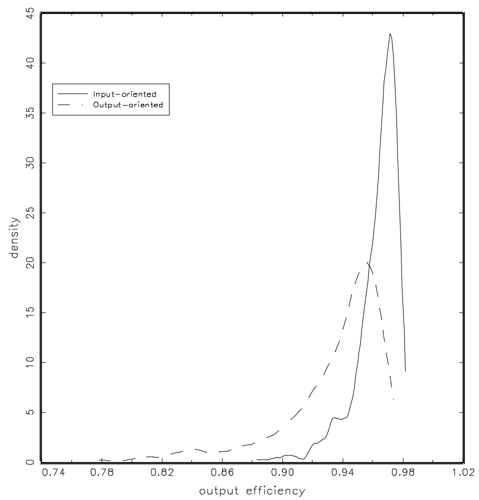
6d. Returns to scale (exponential, prior I)



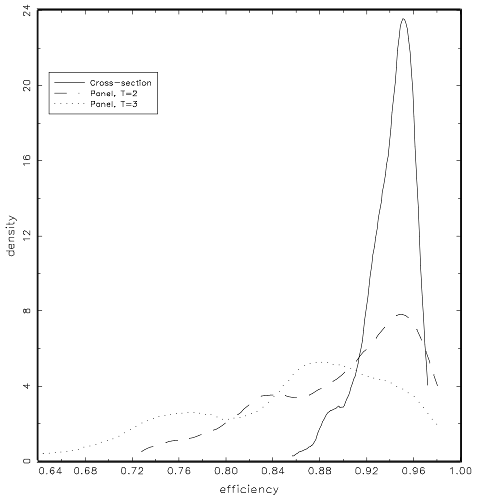
6e. Technical change (exponential, prior I)



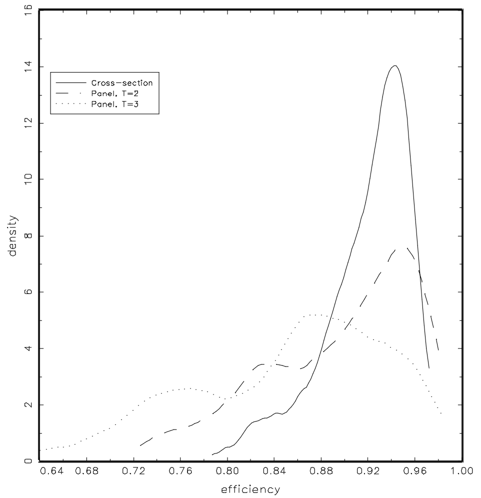
6f. Output efficiency (exponential, prior I)



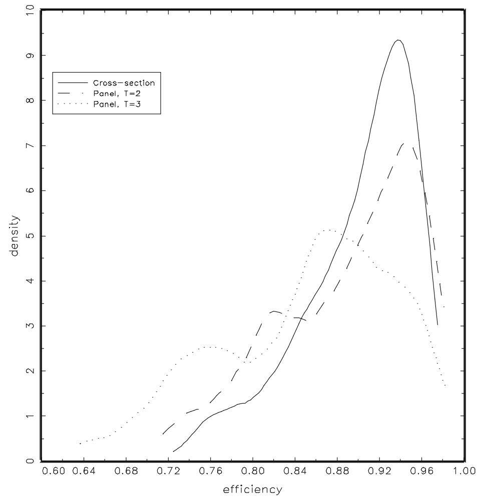
7a. Input-oriented efficiency, half-normal, prior I



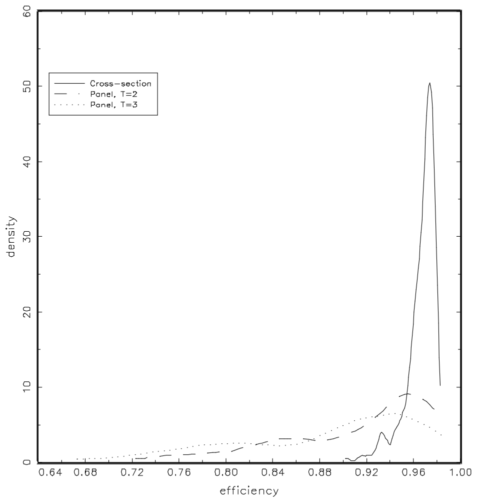
7b. Input-oriented efficiency, half-normal, prior II



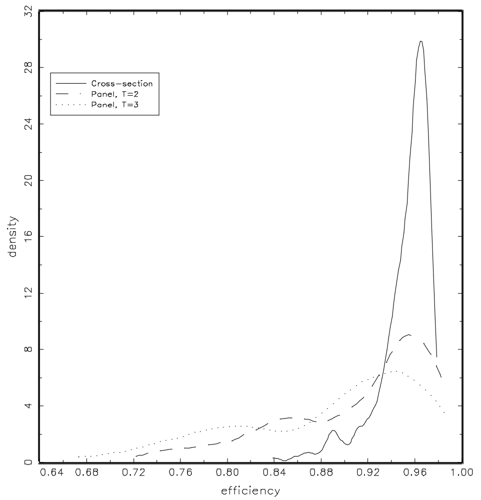
7c. Input-oriented efficiency, half-normal, prior III



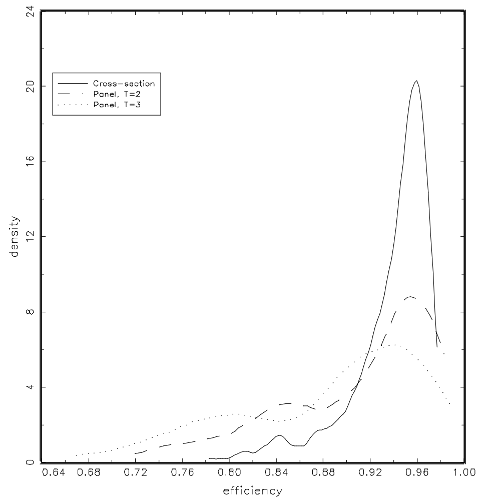
7d. Input-oriented efficiency, exponential, prior I



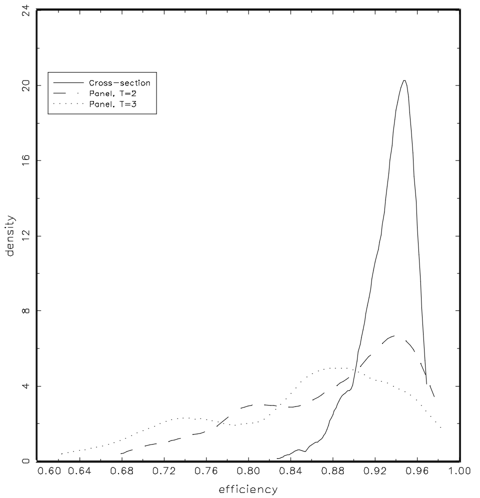
7e. Input-oriented efficiency, exponential, prior II



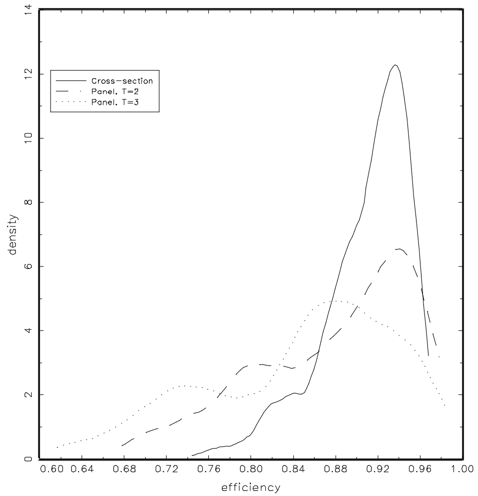
7f. Input-oriented efficiency, exponential, prior III



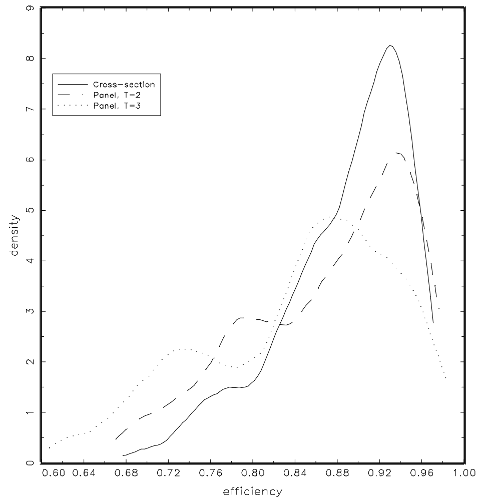
8a. Output-oriented efficiency, half-normal, prior I



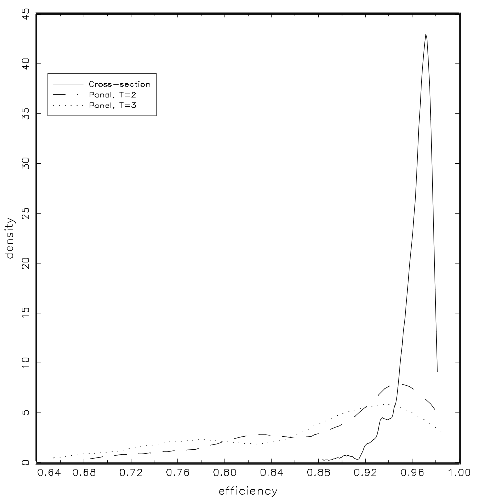
8b. Output-oriented efficiency, half-normal, prior II



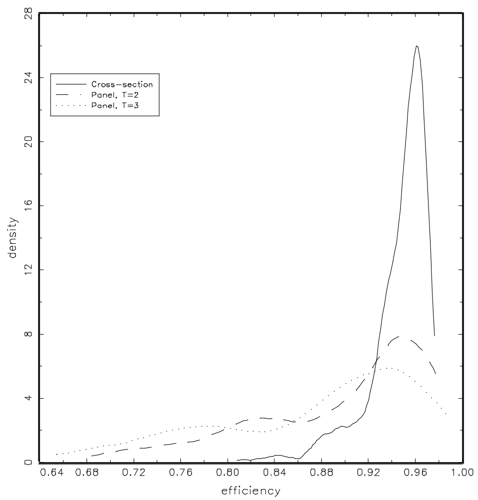
8c. Output-oriented efficiency, half-normal, prior I



8d. Output-oriented efficiency, exponential, prior I



8e. Output-oriented efficiency, exponential, prior II



8f. Output-oriented efficiency, exponential, prior I