Journal of Productivity Analysis, 21, 277–296, 2004

Estimation of Technical Inefficiencies with Heterogeneous Technologies

HO-CHUAN (RIVER) HUANG river@mail.tku.edu.tw
Department of Banking and Finance, Tamkang University, Tamsui, Taipei, Taiwan 251

Abstract

This paper considers the measurement of firm's specific (in)efficiency while allows for the possible heterogeneous technologies adopted by different firms. A flexible stochastic frontier model with random coefficients is proposed to distinguish technical inefficiency from technological differences across firms. Posterior inference of the model is made possible via the simulation-based approach, namely, Markov chain Monte Carlo method. The model is applied to a real data set which has also been considered in Christensen and Greene (1976), Greene (1990), Tsionas (2002), among others. Empirical results show that the regression coefficients can vary across firms, indicating the adoption of heterogeneous technologies by different firms. More importantly, we find that, without considering this possible heterogeneity, the inefficiency of firms can be over-estimated.

JEL Classification: C11, C21, D20

Keywords: stochastic frontier, random-coefficient, Gibbs sampler, Metropolis–Hastings

1. Introduction

There has been much interest, in many areas of economics and finance, in the estimation of production technology and the measurement of a firm's inefficiency relative to this technology. For this purpose, the techniques used in the literature differ mainly by how the difference between inefficiency and random error is distinguished, by how the functional forms are specified, and by how the distributional assumptions are made. Among which, the stochastic frontier model, first developed by Meeusen and van den Broeck (1977) and Aigner et al. (1977), has been one of the most commonly used tools. The usual assumption for the measurement error v_i is i.i.d. normal and independent of the inefficiency u_i. In contrast to the symmetric assumption, the inefficiency term u_i, denoting deviations of the individual units from the frontier, is commonly assumed to follow a one-sided distribution, for example, exponential, truncated (or, half) normal, or gamma.

Based on the distributions, the "composed-error" model can be estimated by maximum likelihood approach. Due to the complexity of the log likelihood function, for example, the normal-gamma model of Greene (1990), however, the classical method often requires cumbersome computation in order to obtain maximum likelihood estimates. Alternatively, the use of Markov chain Monte Carlo (MCMC) method, especially Gibbs sampler with data augmentation, for making Bayesian posterior inference in stochastic frontier models has also been shown to be computationally feasible and conceptually simple. Of particular interest is that implementation of the Bayesian approaches allows us to calculate point estimates and standard deviations of any feature of interest including the inefficiency measure u_i. Methods along this line include Koop et al. (1994, 1995), and Fernandez, et al. (2000, 2002). An updated review on the issue can be found in Koop and Steel (2001).

So far, it is commonly, implicitly or explicitly, adopted in the literature that all firms are assumed to share exactly the same production possibilities and differ only with respect to their degree of inefficiency. In practice, however, firms may adopt different technologies for a variety of reasons. As argued in Tsionas (2002), adoption of a new technology is costly, and firms adopt new technologies only with considerable lags. If costs related to installation and personnel training differ across firms, it follows that at any given point in time there will be some variability in the types of technology used by firms. As a result, we might expect the production possibilities to be different in a cross-section of firms. The later situation is considered in Kalirajan and Obwona (1994) who specify a random coefficient average production function and measure inefficiency using the residuals from a frontier derived by using the mean response coefficients. Unfortunately, their measurement of inefficiency mixes the effects of technological differences and firm-specific inefficiency. In contrast, Tsionas (2002) proposes a stochastic frontier model with random coefficients to separate technical inefficiency from technological differences across firms. Based on Bayesian technique, the inference is performed via the Gibbs sampler with data augmentation algorithm. Although Tsionas' (2002) methodology can avoid the confusion between technological differences and technology-specific inefficiency, the modeling can be extended in at least two directions. First, the exponential distribution assumption on the inefficiency measure u_i can be too restrictive. To relax this assumption, we will postulate that the inefficiency term u_i follows a gamma distribution with the shape parameter not necessarily being an integer. Second, the specification in Tsionas (2002) requires that the regression parameters are all random at same time. Our modeling strategy has the flexibility to allow for only a subset of the parameters to be random while the others to remain fixed.[1]

This paper is organized as follows. Section 2 reviews the basic stochastic frontier modeling strategy. Section 3 presents our flexible random-coefficient stochastic frontier models. Section 4 introduces the MCMC approach, especially the Metropolis–Hastings and Gibbs sampler with data augmentation algorithm. Section 5 derives the full conditional densities for implementation of the MCMC methods. Section 6 applies the novel model to a real data set and discusses the empirical results. Conclusions are finally given in Section 7.

2. Stochastic Frontier Models

2.1. *Basic Modeling*

We start with the basic stochastic frontier framework,

$$y_i = \alpha + x_i'\beta + v_i - u_i, \tag{1}$$

where y_i is the (natural) logarithm of the observed output for the ith firm, α is a non-random scalar intercept term, $x_i = (x_{i1}, x_{i2}, \ldots, x_{ik})'$ is a $k \times 1$ vector of logarithms of inputs, $\beta = (\beta_1, \beta_2, \ldots, \beta_k)'$ is the corresponding $k \times 1$ vector of parameters. The measurement error v_i is assumed to be distributed as i.i.d. $\mathcal{N}(0, \sigma^2)$ and u_i is a non-negative error term indicating the extent of technical inefficiency. To see why, note that the technical efficiency of the ith firm can be defined as $r_i = \exp(-u_i)$ due to our logarithmic specification. Thus, the technical inefficiency of the ith firm is measured by $1 - r_i$. For small values of u_i, the inefficiency $1 - r_i = 1 - \exp(-u_i)$ can be approximated by u_i, so that it is common to use u_i as a measure of technical inefficiency.

Equation (1) can be reformulated as,

$$y_i = \alpha_i + x_i'\beta + v_i, \tag{2}$$

where $\alpha_i = \alpha - u_i$. With this definition, the conventional stochastic frontier model with composite errors can be re-interpreted as a random coefficient (for the intercept term only) framework.

2.2. *The Random-Coefficient Stochastic Frontier Model*

In contrast, Tsionas (2002) argues that, in practice, firms' technologies may be heterogeneous rather than homogeneous. Thus, it would be more appropriate to distinguish technological differences and technology-specific inefficiency rather than simply assume that firms share the same technology. To account for both features, Tsionas (2002) proposes a stochastic frontier production regression to relax the assumption of common frontiers to all firms by allowing the slope coefficients to be random,[2] that is,

$$y_i = \alpha + x_i'\beta_i + v_i - u_i, \tag{3}$$

where the coefficients of the explanatory variables, $\beta_i = (\beta_{i1}, \beta_{i2}, \ldots, \beta_{ik})'$, are assumed to vary across firms so that firm-specific technical inefficiency (u_i) can be separated from technological differentials which are captured by the different values of slope coefficients (β_i).

Specifically, Tsionas (2002) assumes that the individual parameter vectors, $\beta_i, i = 1, 2, \ldots, n$, are independent and follow a k-variate normal distribution,

$$\beta_i \sim \mathcal{N}_k(\beta, \Omega), \tag{4}$$

where β is the $k \times 1$ mean vector and Ω is the $k \times k$ variance-covariance matrix. Note that $\mathcal{N}_k$ represents a normal distribution with dimension k.[3]

Substituting (4) into (3), we can obtain,

$$y_i = \alpha + x_i'\beta + \varepsilon_i - u_i, \tag{5}$$

where ε_i is independently distributed as $\mathcal{N}(0, \sigma^2 + x_i'\Omega x_i)$. Therefore, Tsionas (2002) interprets that the assumption of random coefficients in equation (3) implies a stochastic frontier with heteroskedastic measurement error as shown in equation (5).

2.3. The Distribution of u_i

So far, we have been silent about the specification of the inefficiency term u_i. Conventionally, a distributional assumption is often made for the one-sided disturbance term u_i, for example, exponential or truncated normal.

As an alternative, Greene (1990) proposes a gamma distribution for measuring the technical inefficiency u_i. In particular, given the shape parameter P and scale parameter θ, the probability density function is given by,

$$u_i \sim \mathcal{G}(P, \theta) = \frac{\theta^P}{\Gamma(P)} u_i^{P-1} \exp(-\theta u_i), \tag{6}$$

where $\Gamma(\cdot)$ is the gamma function. This specification nests the commonly used exponential distribution as a special case since the exponential distribution can be obtained by restricting the shape parameter of the gamma distribution to be 1, i.e., $P = 1$. Although this normal-gamma model exhibits richer and more flexible parameterization of the inefficiency distribution than previous ones, the estimation of such model is of only limited success due to complexity of evaluating the log likelihood function. Thus, Greene (2003) proposes an alternative approach to estimate the model by the simulated maximum likelihood estimation. As argued in Greene (2003), the simulation method has proved as a useful tool and seems to work quite well.

Another simulation-based approach is the recent advance in Bayesian analysis, in particular, the MCMC technique. For instance, Koop et al. (1995) consider (three) different specifications of u_i, corresponding to the gamma distribution with different integer shape parameters $P = 1, 2$ or 3. They use the Gibbs sampler, a special case of the MCMC approach, with data augmentation algorithm to make posterior inferences in the stochastic frontier model. However, they restrict their attention to the Erlang case (gamma distribution with integer values of shape parameter), as mentioned above. In contrast, Tsionas (2000) recognizes the constraint and proposes a computational method to free up the restriction that the shape parameter can take only integer values. He also uses the Gibbs sampling approach to make Bayesian inference in the stochastic frontier model with gamma distributed inefficiency terms, where the shape parameter of the gamma distribution is not restricted to be integer values (the Erlang distribution).

3. Our Modeling Strategy

In this paper, we propose an alternative random-coefficient stochastic frontier model which includes many previously analyzed ones as special cases.[4] Particularly, the model is specified as,

$$y_i = \alpha + x_i'\beta + z_i'\gamma_i + v_i - u_i, \tag{7}$$

where z_i is a $k' \times 1 (k' \leq k)$ vector of variables that are a subset of x_i. The corresponding $k' \times 1$ vector of coefficients $\gamma_i = (\gamma_{i1}, \gamma_{i2}, \ldots, \gamma_{ik'})'$ for the ith firm is assumed to be independently, identically and normally distributed with mean vector 0 and variance-covariance matrix Ω, that is,

$$\gamma_i \sim \mathcal{N}_{k'}(0, \Omega). \tag{8}$$

Following Greene (1990, 2003), Koop et al. (1995) and Tsionas (2000), we also assume the gamma distribution as in (6) to capture the technological inefficiency u_i. However, we do not restrict ourselves only to the integer-value shape parameter case. Instead, we consider the non-Erlang case as well, that is, gamma distribution with non-integer shape parameter. As argued in Tsionas (2000), computation of the likelihood function requires the evaluation of integrals with neither closed form solution nor polynomial approximation. Thus, Tsionas (2000) proposes a feasible method to make posterior inference in such model with arbitrary values for the shape parameter.

As claimed earlier, our model is rather flexible and includes many popular stochastic frontier approaches as special cases. In particular, we discuss some cases below. When z_i is a null set and $P = 1$, model in (7) becomes the one considered in Aigner et al. (1977) and Meeusen and van den Broeck (1977), that is, the exponential case. In contrast, if z_i is an empty set and $P = 1, 2, 3$, our model becomes the one investigated in Koop et al. (1995), that is, the Erlang case. The case discussed in Tsionas (2000) can also nested by restricting z_i to be a null set and freeing up the shape parameter P to be non-integer. In case when z_i is identical to x_i, equation (7) appears to be the random-coefficient stochastic frontier model in equation (3) discussed by Tsionas (2002). The equivalence can be seen by setting $P = 1$ and β_i in equation (3) to be $\beta + \gamma_i$ in equation (7). If we do not consider the one-sided disturbance term u_i, the restriction of $z_i = x_i$ makes our model identical to the one proposed by Kalirajan and Obwona (1994). Of course, z_i need not be a null set or identical to x_i. It can be a subset of x_i, meaning that only some of the coefficients are random while others remain constant.

To fix the idea, the joint posterior distribution of parameters $(\alpha, \beta, \sigma^2, \gamma, \Omega, P, \theta, u)$ based on the observed data $y = (y_1, y_2, \ldots, y_n)'$, $X = (x_1', x_2', \ldots, x_n')'$ and $Z = (z_1', z_2', \ldots, z_n')'$ can be expressed as a product of the prior $\pi(\alpha, \beta, \sigma^2, \Omega, P, \theta)$ and the likelihood function,[5]

$$\pi(\alpha, \beta, \sigma^2, \gamma, \Omega, P, \theta, u \mid y, X, Z) = \pi(\alpha, \beta, \sigma^2, \Omega, P, \theta)$$
$$\times \prod_{i=1}^{n} \left\{ (2\pi\sigma^2)^{-1/2} \exp\left[-\frac{1}{2\sigma^2} (y_i + u_i - \alpha - x_i'\beta - z_i'\gamma_i)^2 \right] \right\}$$
$$\times \prod_{i=1}^{n} \left\{ \frac{\theta^P}{\Gamma(P)} u_i^{P-1} \exp(-\theta u_i) \right\} \times \prod_{i=1}^{n} \left\{ (2\pi)^{-k'/2} |\Omega|^{-1/2} \exp\left[-\frac{1}{2} \gamma_i' \Omega^{-1} \gamma_i \right] \right\}, \qquad (9)$$

where $\gamma = (\gamma_1', \gamma_2', \ldots, \gamma_n')'$ and $u = (u_1, u_2, \ldots, u_n)'$.

Note that the posterior of the "structural" parameters, namely $(\alpha, \beta, \sigma^2, \gamma, \Omega, P, \theta)$, can be obtained by integrating out the nuisance parameter u, that is,

$$\pi(\alpha, \beta, \sigma^2, \gamma, \Omega, P, \theta \mid y, X, Z) = \int \pi(\alpha, \beta, \sigma^2, \gamma, \Omega, P, \theta, u \mid y, X, Z) du. \qquad (10)$$

Although equation (10) seems awkward, the Bayesian approach outlined later on, via Gibbs sampler or Metropolis–Hastings with data augmentation algorithm, allows us to obtain exact finite sample properties of all features of interest (including firm-specific inefficiency) and surmount some difficult statistical problems involved with classical estimation.

4. Bayesian Inference

In order to conduct Bayesian posterior inference, we need to specify the prior distribution on the parameters $(\alpha, \beta, \sigma^2, \Omega, P, \theta)$. It is worthwhile noting that the use of improper priors in stochastic frontier models can cause problems such as lack of existence of the posterior itself, for example, Ritter and Simar (1997) and Fernandez et al. (1997). As a result, the values for the hyperparameters are chosen so that they imply relatively vague but proper priors. In particular, we assume the priors are:

$$\pi(\alpha, \beta, \sigma^2, \Omega, P, \theta) = \pi(\alpha)\pi(\beta)\pi(\sigma^2)\pi(\Omega)\pi(P)\pi(\theta). \qquad (11)$$

Specifically, the prior for the constant term α is a univariate normal distribution with mean α_0 and variance A_0, that is,

$$\alpha \sim \mathcal{N}(\alpha_0, A_0), \qquad (12)$$

and the prior assumed on the frontier parameters β takes the following form,

$$\beta \sim \mathcal{N}_k(\beta_0, B_0), \qquad (13)$$

that is, a k-dimensional normal distribution with mean vector β_0 and variance-covariance matrix B_0. In the following empirical application, we specify $\alpha_0 = 0$, $A_0 = 100, \beta_0 = (0,0,0,0)'$ and $B_0 = 100 \times I_4$. In general, those values represent relatively vague priors. Alternatively, we could set the priors for α and β to be flat as does in Tsionas (2002), reflecting no prior information about parameter means.

We define the prior for σ^2 through a gamma distribution on the precision σ^{-2},

$$\sigma^{-2} \sim \mathcal{G}\left(\frac{v_0}{2}, \frac{\delta_0}{2}\right). \tag{14}$$

In many applications, the hyperparameters v_0 and δ_0 are set to be zeros to have a standard diffuse Jeffrey's prior. However, as shown in Fernandez et al. (1997), a proper prior is required in order to make the posterior distribution well defined in the stochastic frontier framework. As a result, we set v_0 to be 0 but let δ_0 be 0.01.

Similarly, the prior for the $k' \times k'$ variance-covariance matrix Ω is defined through a Wishart distribution on the precision matrix Ω^{-1}, expressed as,

$$\Omega^{-1} \sim \mathcal{W}_{k'}(r_0, R_0). \tag{15}$$

Jeffrey prior can be obtained by restricting $r_0 = 0$ and $R_0 = 0$. However, we set $r_0 = 0$ and $R_0 = 100 \times I_4$ to have a relatively vague but proper prior.

Finally, the priors for the shape parameter P and the scale parameter θ are chosen as,

$$P \sim \mathcal{G}(a_0, b_0), \tag{16}$$

$$\theta \sim \mathcal{G}(c_0, d_0), \tag{17}$$

i.e., both priors are gamma distributed. In Tsionas (2002), he assumes that inefficiency is exponentially distributed with prior on the parameter θ being exponential. In terms of our notations, this means that $P = 1$, $c_0 = 1$ and $d_0 = -\ln(r^*)$ where r^* is the prior median efficiency as introduced by van den Broeck et al. (1994). Similarly, we follow van den Broeck et al. (1994) and Tsionas (2002) to specify $c_0 = 1$ and $d_0 = -\ln(r^*)$. In addition, we assume that $a_0 = 1$ and $b_0 = 1$ such that the prior for P is exponential with mean being equal to 1 as in Tsionas (2002).[6]

Given the specified priors, estimation can be carried out via the MCMC approach, in particular, the Gibbs sampler and Metropolis–Hastings algorithm with data augmentation. Appendix A provides a brief review of the methods.[7] For the relevant full conditional densities needed for implementation of the MCMC approach, readers are referred to Appendix B for more details.

5. Empirical Applications

The data set used to illustrate the technique is collected by Christensen and Greene (1976) for a total of 123 electric utility companies in the United States. The same data set has been previously analyzed by Greene (1990), van den Broeck et al. (1994), Koop et al. (1995) and Tsionas (2002). For comparison purpose, we first estimate the fixed-coefficient Cobb–Douglas cost function specified as,

$$\ln\left(\frac{c}{p_f}\right)_i = \alpha + \beta_1 \ln\left(\frac{p_l}{p_f}\right)_i + \beta_2 \ln\left(\frac{p_k}{p_f}\right)_i + \beta_3 \ln q_i + \beta_4 \ln^2 q_i + v_i - u_i, \tag{18}$$

where c is total cost, q is output, and p_l, p_k and p_f are the unit prices of labor, capital and fuel, respectively. The symmetric disturbance term v_i is i.i.d. $\mathcal{N}(0, \sigma^2)$ and the non-negative error u_i is i.i.d. $\mathcal{G}(P, \theta)$.

In contrast, the random-coefficient counterpart takes the following form,

$$\begin{aligned}\ln\left(\frac{c}{p_f}\right)_i &= \alpha + \beta_1 \ln\left(\frac{p_l}{p_f}\right)_i + \beta_2 \ln\left(\frac{p_k}{p_f}\right)_i + \beta_3 \ln q_i + \beta_4 \ln^2 q_i \\ &\quad + \gamma_{i1} \ln\left(\frac{p_l}{p_f}\right)_i + \gamma_{i2} \ln\left(\frac{p_k}{p_f}\right)_i + \gamma_{i3} \ln q_i + \gamma_{i4} \ln^2 q_i + v_i - u_i,\end{aligned} \tag{19}$$

or, more compactly,

$$y_i = \alpha + x_i'\beta + z_i'\gamma_i + v_i - u_i, \tag{20}$$

where $y_i = \ln(c/p_f)_i$, $x_i = (\ln(p_l/p_f)_i,\ \ln(p_k/p_f)_i,\ \ln q_i, \ln^2 q_i)'$, $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)'$, $z_i = x_i$, and $\gamma_i = (\gamma_{i1}, \gamma_{i2}, \gamma_{i3}, \gamma_{i4})'$.

Both models are estimated via the Metropolis–Hastings and Gibbs sampler with data augmentation algorithm by assuming relatively diffuse priors. The chain is then run for 10,000 iterations. We collect the last 5,000 sample variates after discarding the first 5,000 draws. As a result, the following results are based on 5,000 Gibbs output for making posterior inference. Note that we have used the Metropolis–Hastings algorithm in two places, one for u_i and the other for P. We choose the exponential distribution as the proposal density in both cases since the values must be non-negative for both parameters. In general, the acceptance rates for generating u_i are around 0.25 to 0.30 while the acceptance rate for generating P is approximately between 0.40 and 0.45.

The left panel of Table 1 reports the posterior moments of the fixed-coefficient stochastic frontier model and Figure 1 shows the boxplots of the relevant parameters. First, note that the posterior means of β coefficients are all positive as expected despite that the coefficient of the squared output term, $\ln^2 y_i$, is not significantly different from zero, judged by the 95% Bayesian confidence interval $[-0.0512, 0.1739]$. In addition, the sum of the β estimates equals to 0.7667, implying that the technology is decreasing returns to scale. The estimate of the variance term σ^2 is very small with a value 0.0141. The shape and scale parameters, i.e., P and θ, of the gamma distribution of u_i have posterior means 0.9575 and 9.9025, respectively. These results are comparable to those found in Koop et al. (1995) and Tsionas (2002). Note that the posterior mean of the shape parameter P is very close to 1, justifying the use of the exponential distribution for the inefficiency term u_i as in Tsionas (2002). As discussed earlier, our main concern is on the measurement of firm-specific efficiency. Figure 2 presents the kernel density of the (mean) efficiency measures of all firms. It is apparent that the efficiency distribution is highly left-

Table 1. Fixed- vs. random-coefficient model.

	Fixed-coefficient		Random-coefficient	
	Mean	Std	Mean	Std
α	− 7.4784	0.3146	− 7.2171	0.5843
β_1	0.4447	0.0390	0.3668	0.0726
β_2	0.0284	0.0027	0.0335	0.0055
β_3	0.2346	0.0626	0.2517	0.1101
β_4	0.0590	0.0575	0.0695	0.1076
σ^2	0.0124	0.0027	0.0014	0.0012
P	0.9575	0.1136	0.9063	0.2252
θ	9.9025	2.2668	77.4337	16.5072
ω_{11}			0.0222	0.0239
ω_{21}			− 0.0014	0.0016
ω_{22}			0.0001	0.0001
ω_{31}			− 0.0156	0.0166
ω_{32}			0.0008	0.0011
ω_{33}			0.0131	0.0121
ω_{41}			− 0.0010	0.0142
ω_{42}			− 0.0001	0.0010
ω_{43}			0.0002	0.0103
ω_{44}			0.0153	0.0179

Note: The posterior means and posterior standard deviations are obtained using 5,000 simulated draws after discarding the first 5,000 variates to mitigate the start-up effect.

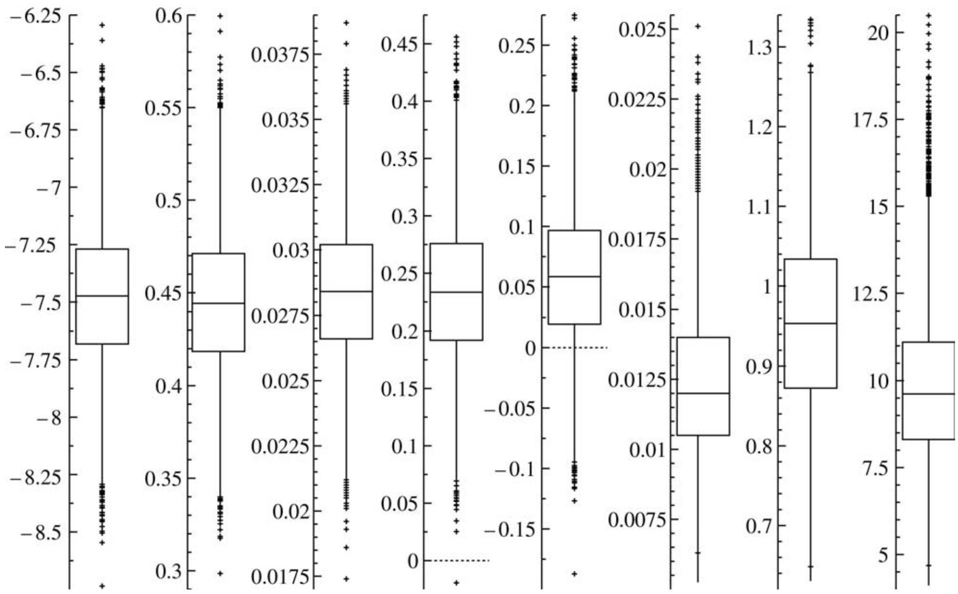

Figure 1. The boxplots of the fixed-coefficient model: $\alpha, \beta_1, \beta_2, \beta_3, \beta_4, \sigma^2, P, \theta$ (left to right).

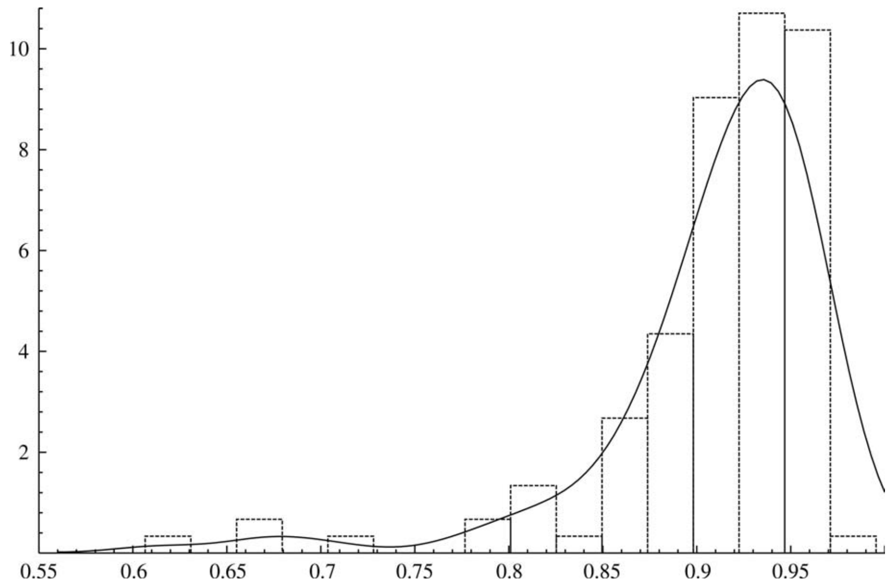

Figure 2. The kernel density of mean efficiencies obtained from the fixed-coefficient stochastic frontier model.

skewed and exhibits large variation over firms. In addition, the average efficiency measure of those firms has a value 0.9103, which is very close to that of the representative firm considered in Koop et al. (1995).

In contrast, we also report the posterior results of the random-coefficient model in the right panel of Table 1. The posterior distributions in terms of boxplots are also presented in Figure 3. The posterior means of the common coefficients β, again, are all positive but with larger variation according to the corresponding (larger) posterior standard deviations. In fact, the posterior means of $\beta_1, \beta_2, \beta_3$ and β_4 are 0.3668, 0.0335, 0.2517 and 0.0695 with posterior standard deviations 0.0726, 0.0055, 0.1101, 0.1076, respectively. In addition, we calculate the posterior means of the random-coefficient vector $\gamma_i = (\gamma_{i1}, \gamma_{i2}, \gamma_{i,3}, \gamma_{i4})'$ to be $-0.0005, 0.0000, 0.0007$ and 0.0004, respectively. By taking the sum of the mean values of both common and random coefficients, we can obtain the mean coefficients to be 0.3663, 0.0335, 0.2524 and 0.0699, respectively. Compared to the estimated values in the fixed-coefficient model, β_1 is smaller but β_2, β_3 and β_4 are all larger in the random-coefficient model. The variance term σ^2 has a posterior mean 0.0014 which is almost one-eighth of the corresponding estimate in the fixed-coefficient model. The mean shape estimate of P is 0.9063, slightly less than 0.9575 as in the fixed-coefficient case. The largest difference between these two models pertains to the posterior mean of the scale parameter θ which carries important implications for measuring efficiencies. In the random-coefficient stochastic model, the posterior mean of θ is 77.4337 with posterior standard deviation 16.5072. The value is about eight times larger compared to the estimate in the fixed-coefficient stochastic frontier model. Even though, the

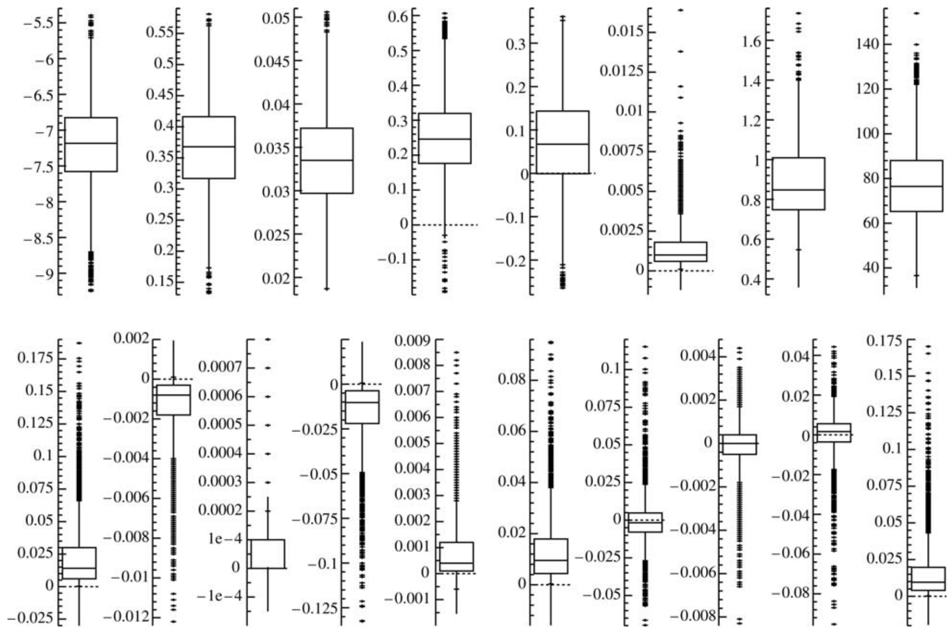

Figure 3. The boxplots of the random-coefficient model: $\alpha, \beta_1, \beta_2, \beta_3, \beta_4, \sigma^2, P, \theta$ (left to right, top panel); $\omega_{11}, \omega_{21}, \omega_{22}, \omega_{31}, \omega_{32}, \omega_{33}, \omega_{41}, \omega_{42}, \omega_{43}, \omega_{44}$ (left to right, bottom panel).

estimate is comparable to that found in Tsionas (2002). By assuming an exponential distribution for the inefficiency term u_i, that is, gamma distribution with shape parameter $P = 1$, Tsionas (2002) finds a posterior mean value for θ to be 71.77. Given those estimates, we are able to calculate the mean efficiency of all firms. Surprisingly, the mean efficiency increases dramatically to 0.9891. Figure 4 shows the kernel density of the mean efficiencies of all firms in the random-coefficient stochastic frontier model. In contrast to the corresponding kernel density in the fixed-coefficient model, those efficiency measures are centered around the mean value. Besides, the graph also shows much less variation over firms.

For illustrative purpose, we also present the posterior mean efficiencies (with posterior standard deviations) of the first five firms from both the fixed- as well as random-coefficient stochastic frontier models in Table 2. It is worth stressing that the posterior means (standard deviations) of efficiency measures for all five firms are higher (smaller) in the random-coefficient model than those obtained from the fixed-coefficient case. Figure 5 presents the kernel densities of efficiency measures for those five firms. It is very impressive that the efficiency measures of those five firms are very close to 1, that is, fully efficient. This observation is in accord to that found in Tsionas (2002). Specifically, Tsionas (2002) argues that "the random coefficient model implies that full efficiency is about seven times more likely compared to the fixed coefficients stochastic frontier model." In addition, we find that the difference of efficiency measures can be quite large under fixed-coefficient and random-coefficient

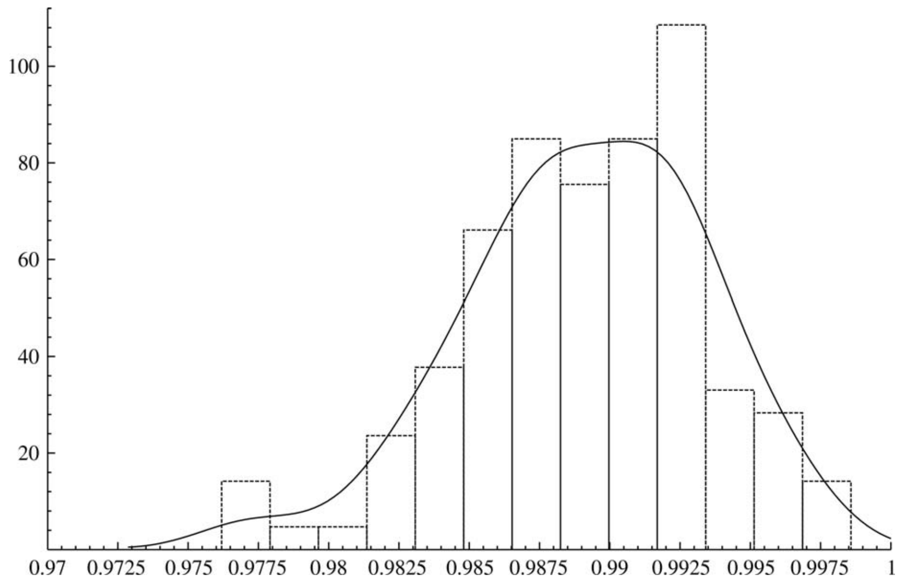

Figure 4. The kernel density of mean efficiencies obtained from the random-coefficient stochastic frontier model.

models. For instance, the mean efficiency of the first firm is only 0.6714 obtained from the fixed-coefficient model while it dramatically increases to 0.9953 when a random-coefficient stochastic frontier regression is considered. This evidence indicates the measured efficiency can be higher once we allow for firms to adopt different technologies.

Although we believe that allowing for heterogeneous technologies makes more sense and is much closer to real world, a formal comparison can be made via the Bayes factor, for example, Chib (1995) and Chib and Jeliazkov (2001).[8] Let $m(y)$ denote the marginal likelihood, $f(y\,|\,\theta)$ the likelihood function, $\pi(\theta)$ the prior density, and $\pi(\theta\,|\,y)$ the posterior density, then the computationally convenient log

Table 2. Posterior moments of efficiency measures.

Firm	1	2	3	4	5
	Fixed-coefficient Model				
Mean	0.6714	0.9700	0.9448	0.8911	0.9479
Std	0.1007	0.0222	0.0414	0.0605	0.0306
	Random-coefficient Model				
Mean	0.9953	0.9912	0.9861	0.9923	0.9926
Std	0.0059	0.0086	0.0134	0.0112	0.0060

Note: The results are based on 5,000 Gibbs draws after discarding the first 5,000 sample variates.

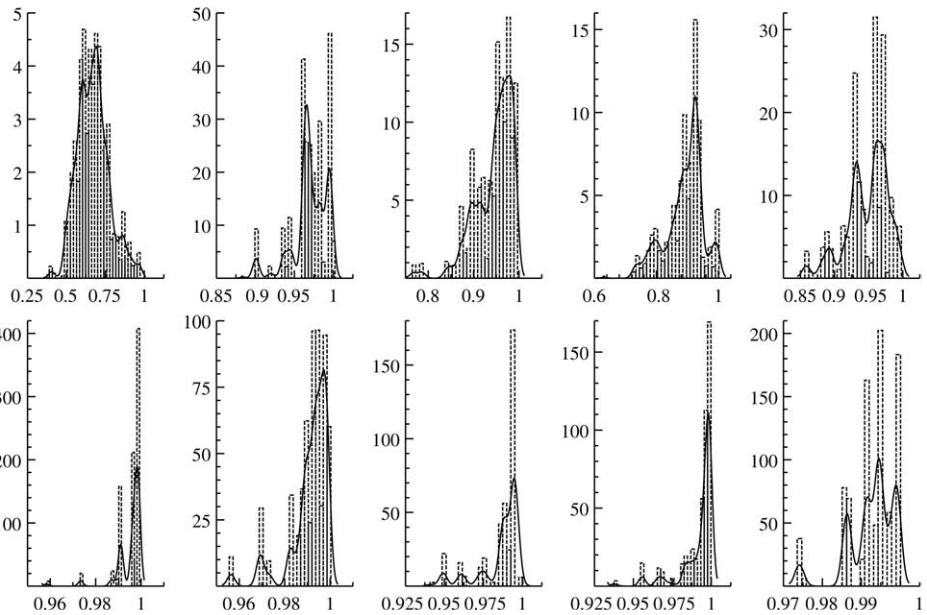

Figure 5. The kernel densities of the efficiencies of the representative (the first five, left to right) firms obtained from fixed-coefficient (top panel) as well as random-coefficient (bottom panel) stochastic frontier models.

marginal likelihood can be obtained by,

$$\ln \hat{m}(y) = \ln f(y \mid \theta^*) + \ln \pi(\theta^*) - \ln \hat{\pi}(\theta^* \mid y),$$

where θ^* is often chosen to be a high density point and $\hat{\pi}(\theta^* \mid y)$ can be estimated by the algorithm discussed in Chib (1995) if all integrating constants of the full conditional distributions in the Gibbs sampler are known, or/and Chib and Jeliazkov (2001) if some of full conditionals are intractable, for example, the full conditional density of P in our case. As a result, Bayes factor for any two models k and h, that is, $m(y \mid M_k)/m(y \mid M_h)$, can be calculated by,

$$\hat{B}_{kh} = \exp\{\ln \hat{m}(y \mid M_k) - \ln \hat{m}(y \mid M_h)\}.$$

The Bayes factor for the fixed-coefficient model versus the random-parameter model is 4.1239464×10^{-6}, suggesting strong evidence in favor of the random-parameter stochastic frontier model. In addition, we also provide the kernel densities of the random coefficients for three representative firms $i = 1, 3$, and 5 in Figure 6. It can be seen that the densities exhibit more or less variations, indicating that different firms may adopt different technologies. The findings of the Bayes factor and the graph convey an important message to empirical researchers working with stochastic frontier models, namely, more appropriate measure of technical (in)efficiency has to take into account the possible heterogeneity of technologies over firms.

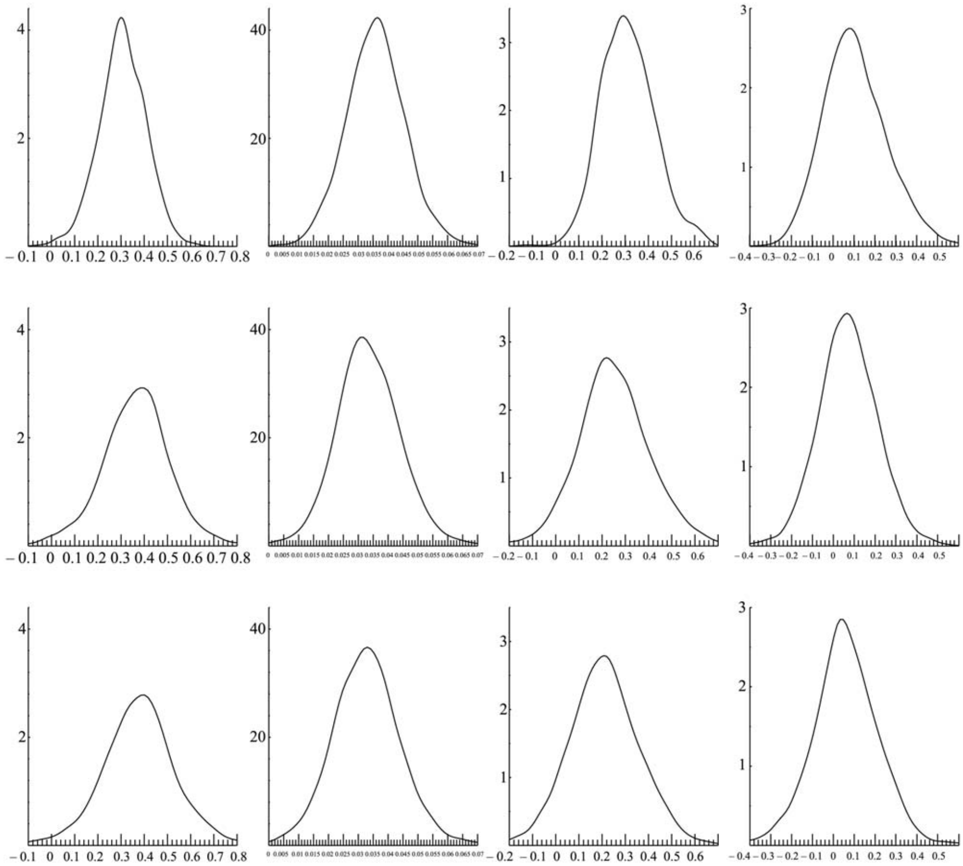

Figure 6. The kernel densities of the parameters (from left to right, β_{i1}, β_{i2}, β_{i3} and β_{i4}) for the representative firms (from top to bottom, the *i*th firm, $i = 1, 3$ and 5). Note that $\beta_{ij} = \beta_j + \gamma_{ij}$.

6. Conclusions

This paper considers the measurement of firm's specific (in)efficiency while allows for the possible heterogeneous technologies adopted by different firms. A very flexible stochastic frontier model with random coefficients is proposed to distinguish technical inefficiency from technological differences across firms. Posterior inference of the model is made possible via the simulation-based approach, namely, MCMC method. In particular, for those parameters which have full conditionals with standard form are directly simulated using Gibbs sampler with data augmentation algorithm while for those parameters with non-standard full conditionals are drawn using the Metropolis–Hastings algorithm. The model is applied to a real data set which has also been considered in Christensen and Greene (1976), Greene (1990), Tsionas (2002), among others. Empirical results show that the regression coefficients can vary across firms, indicating the adoption of heterogeneous technologies by different firms. More importantly, without considering this possible heterogeneity, the inefficiency of firms can be over-estimated. As a result, we believe that the novel techniques proposed in this paper might allow for better understanding of firm efficiency than do traditional methods.

Appendix A: The MCMC Approach

Recently, statisticians and econometricians have been increasingly relied on MCMC methods to simulate complicated, non-standard multivariate distributions. Among which, the Gibbs sampler, first introduced by Geman and Geman (1984), is an algorithm for generating random variates from an intractable marginal distribution indirectly from the full conditional distributions which are usually available in a simple form.

In many applications, direct sampling from a posterior distribution $\pi(\theta \mid y)$ is commonly infeasible. However, assume that the full conditional distributions of the latent variable y^* along with the partitioned parameters $\theta = (\theta_1, \theta_2, \ldots, \theta_p)'$ are all available and in standard forms. The Gibbs sampler with data augmentation algorithm generates posterior variates $(y^{*(t)}, \theta^{(t)})$ by recursively drawing from the following distributions,

$$\pi\left(y^* \mid y, \theta^{(t-1)}\right), \quad \pi\left(\theta_i \mid y, y^{*(t)}, \theta_j^{(t)}, \theta_k^{(t-1)}\right), \tag{21}$$

where $j\langle i, k\rangle i$ and $i = 1, 2, \ldots, p$. As shown in Gelfand and Smith (1990), the random variate $(y^{*(t)}, \theta^{(t)})$, converges in distribution to that of joint posterior density of the latent variable y^* and the parameter θ as $t \to \infty$. Furthermore, according to the ergodic theorem,

$$\frac{1}{t}\sum_{i=1}^{t} f(\theta^{(i)}) \to \int f(\theta)\pi(\theta)d\theta = E[f(\theta)], \tag{22}$$

i.e., the averages of functions evaluated at sample values (ergodic averages) converge to their expected values under the target density $\pi(\theta)$.

However, in cases where the full conditional distributions do not have a standard form direct draws may be possible. In those cases, the Metropolis–Hastings (MH) algorithm can be tailored to generate the required sample variates. By repeating for $t = 1, 2, \ldots, T$ and specifying an initial value of θ_i, say $\theta_i^{(0)}$, the MH algorithm proceeds by generating a candidate value θ_i' from a candidate-generating density $q(\theta_i^{(t-1)}, \theta_i')$ and v from the uniform density $\mathscr{U}(0, 1)$. Then, we set the next draw $\theta_i^{(t)} = \theta_i'$ if

$$v \leq \alpha\left(\theta_i^{(t-1)}, \theta_i'\right) = \min\left[\frac{\pi\left(\theta_i' | y, y^{*(t)}, \theta_j^{(t)}, \theta_k^{(t-1)}\right) q\left(\theta_i', \theta_i^{(t-1)}\right)}{\pi\left(\theta_i^{(t-1)} | y, y^{*(t)}, \theta_j^{(t)}, \theta_k^{(t-1)}\right) q\left(\theta_i^{(t-1)}, \theta_i'\right)}, 1\right], \tag{23}$$

otherwise, we return the previous draw $\theta_i^{(t-1)}$ as a value from the target distribution. An expository review of the MH algorithm can be found in Chib and Greenberg (1995).

Appendix B: The Full Conditional Distributions

Given available the priors, we are now ready to conduct the Bayesian posterior analysis. The idea is to treat the latent inefficiency term u as an additional parameter to be sampled along with the other parameters. Once the values of u are simulated, the inference for the remaining unknowns becomes straightforward. Thus, we describe in details how to obtain the relevant full conditional distributions for implementation of the Gibbs sampler with data augmentation algorithm.[9]

- Full conditional density of $u_i, i = 1, 2, \dots, n$:
  First, the full conditional density of the augmented parameter, i.e., the latent inefficiency term $u_i, i = 1, 2, \dots, n$, has the following form,

$$u_i|\alpha, \beta, \sigma^2, \gamma, \Omega, P, \theta \propto u_i^{P-1} \exp\left\{-\frac{1}{2\sigma^2}(u_i - w_i)^2 - \theta u_i\right\} I(u_i \geq 0), \tag{24}$$

  where $w_i = \alpha + x_i'\beta + z_i'\gamma_i - y_i$. The density is in a non-standard form which makes direct draws difficult, if not impossible. However, an independence chain Metropolis–Hastings algorithm can be tailored to simulate the adequate values of u_i. The basic idea is to draw values of u_i from a candidate-generating density (user-specified) and we switch to the candidate value with a certain probability as previously described.

- Full conditional density of α:
  From (9) and (12), the full conditional distribution of α can be shown to be,

$$\alpha|\beta, \sigma^2, \gamma, \Omega, P, \theta, u \sim \mathcal{N}(\hat{\alpha}, A), \tag{25}$$

  where

$$\hat{\alpha} = A\left[A_0^{-1}\alpha_0 + \sigma^{-2}\sum_{i=1}^{n}\left(y_i + u_i - x_i'\beta - z_i'\gamma_i\right)\right],$$

  and

$$A = \left[A_0^{-1} + n\sigma^{-2}\right]^{-1}.$$

- Full conditional density of β:
  Similarly, the full conditional density of β can be shown to be,

$$\beta \mid \alpha, \sigma^2, \gamma, \Omega, P, \theta, u \sim \mathcal{N}_k\left(\hat{\beta}, B\right), \tag{26}$$

where

$$\hat{\beta} = B\left[B_0^{-1}\beta_0 + \sigma^{-2}\sum_{i=1}^{n} x_i(y_i + u_i - \alpha - z_i'\gamma_i)\right],$$

and

$$B = \left[B_0^{-1} + \sigma^{-2}\sum_{i=1}^{n} x_i x_i'\right]^{-1}.$$

- Full conditional density of σ^2:
  For simplicity, we will compute the conditional posterior density of σ^{-2}. Given the prior in (14), it can be shown that σ^{-2} is gamma distributed as,

$$\sigma^{-2} \mid \alpha, \beta, \gamma, \Omega, P, \theta, u \sim \mathcal{G}\left[\frac{v_0 + n}{2}, \frac{\delta_0 + \sum_{i=1}^{n}(y_i + u_i - \alpha - x_i'\beta_i - z_i'\gamma_i)^2}{2}\right]. \tag{27}$$

- Full conditional density of $\gamma_i, i = 1, 2, \ldots, n$:
  Following the same derivation procedure, it is shown that the full conditional distribution of $\gamma_i, i = 1, 2, \ldots, n$, is normally distributed as,

$$\gamma_i \mid \alpha, \beta, \sigma^2, \Omega, P, \theta, u \sim \mathcal{N}_{k'}(\hat{\gamma}_i, \Gamma_i), \tag{28}$$

where the mean vector is

$$\hat{\gamma}_i = \Gamma_i[\sigma^{-2}z_i(y_i + u_i - \alpha - x_i'\beta)],$$

and the variance-covariance matrix is

$$\Gamma_i = [\Omega^{-1} + \sigma^{-2}z_i z_i']^{-1}.$$

- Full conditional density of Ω^{-1}:
  With the conjugate Wishart prior for Ω^{-1}, its full conditional distribution is also Wishart distributed as,

$$\Omega^{-1} \mid \alpha, \beta, \sigma^2, \gamma, P, \theta, u \sim \mathcal{W}_{k'}\left\{r_0 + n, \left[R_0^{-1} + \sum_{i=1}^{n}\gamma_i\gamma_i'\right]^{-1}\right\}. \tag{29}$$

- Full conditional density of P:
  The full conditional distribution for the shape parameter is given by

$$P \mid \alpha, \beta, \sigma^2, \gamma, \Omega, \theta, u \propto \frac{\theta^{nP}}{\Gamma(P)^n} P^{a_0 - 1} \exp\left\{ \left(\sum_{i=1}^{n} \ln u_i - b_0 \right) P \right\}. \tag{30}$$

  This distribution is non-standard but the random variates can be obtained in the same manner as for u_i by using the independence chain Metropolis–Hastings algorithm.

- Full conditional density of θ:
  In contrast, the full conditional density of the scale parameter θ is of standard form as,

$$\theta \mid \alpha, \beta, \sigma^2, \gamma, \Omega, P, u \sim \mathcal{G}\left[c_0 + nP, d_0 + \sum_{i=1}^{n} u_i \right]. \tag{31}$$

  Clearly, it is gamma distributed.

Acknowledgments

The author is grateful to NSC for financial support under grant number NSC 89-2415-H-128-011. Helpful discussions and suggestions from Cliff J. Huang, Hung-Jen Wang and two anonymous referees are highly appreciated.

Notes

1. In order to decide whether a coefficient is fixed or random, two referees suggest that we can estimate various combinations of models and compare them using Bayes factor.
2. Although the setup of Tsionas (2002) is in a panel-data framework, it corresponds to cross-section data case when the time period is set to be 1.
3. From a Bayesian perspective, the distribution on the random coefficient can be interpreted as the prior density.
4. Thanks to one referee for pointing out that we might ask too much from the (cross-sectional) data by asking the data to tell us an inefficiency measure for each firm and different regression parameters for each firm. As suggested by one referee, we might want to remind the readers that the models discussed below are really models with panel data with time-invariant technical efficiency and time-invariant parameters.
5. Note that the distribution in equation (8) can be treated as the prior on the parameters γ_i in equation (7) as in a hierarchical model. Note also that we treat the unobserved $u_i, i = 1, 2, \ldots, n$, as additional parameters to be estimated along with other regression parameters.
6. For illustrative purpose, we follow the suggestion of a referee to draw 10,000 variates each time from those priors and calculate the mean efficiency. By repeating the procedure 10,000 times, we find that the mean efficiency is roughly about 0.8 which might be a reasonable approximation.
7. In addition, a referee suggests that model comparison can be implemented by obtaining the Bayes factor. Chib (1995) and Chib and Jeliazkov (2001) provide an easy and feasible method by directly

using the sampled output from Gibbs sampler or Metropolis–Hastings algorithm to calculate Bayes factor in order to compare a collection of models.

8. I thank one referee for suggesting this approach for model comparison.
9. Given the specified proper priors, the existence of posterior distribution as well as posterior moments is proved by Fernandez et al. (1997) and Tsionas (2002).

References

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

Chib, S. (1995). "Marginal Likelihood from the Gibbs Output." *Journal of the American Statistical Association* 90, 1313–1321.

Chib, S. and E. Greenberg. (1995). "Understanding the Metropolis–Hastings Algorithm." *American Statistician* 49, 327–335.

Chib, S. and I. Jeliazkov. (2001). "Marginal Likelihood from the Metropolis–Hastings Output." *Journal of the American Statistical Association* 96, 270–281.

Christensen, L. and W. Greene. (1976). "Economies of Scale in U.S. Electric Power Generation." *Journal of Political Economy* 84, 655–676.

Fernandez, C., G. Koop and M. Steel. (2002). "Multiple-Output Production with Undesirable Outputs: An Application to Nitrogen Surplus in Agriculture." *Journal of the American Statistical Association* 97, 432–442.

Fernandez, C., G. Koop and M. Steel. (2000). "A Bayesian Analysis of Multiple-Output Production Frontiers." *Journal of Econometrics* 98, 47–79.

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, 169–193.

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.

Geman and Geman. (1984). "Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721–741.

Greene, W. H. (2003). "Simulated Likelihood Estimation of the Normal-Gamma Stochastic Frontier Function." in *Journal of Productivity Analysis* 19, 179–190.

Greene, W. H. (1990). "A Gamma-Distributed Stochastic Frontier Model." *Journal of Econometrics* 46, 141–163.

Kalirajan, K. P. and M. B. Obwona. (1994). "Frontier Production Function: The Stochastic Coefficients Approach." *Oxford Bulletin of Economics and Statistics* 56, 87–96.

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. F. J. Steel. (1997). "Bayesian Efficiency Analysis Through Individual Effects: Hospital Cost Frontiers." *Journal of Econometrics* 76, 77–105.

Koop, G., J. Osiewalski and M. F. J. Steel. (1994). "Bayesian Efficiency Analysis with a Flexible Form: The AIM Cost Function." *Journal of Business & Economic Statistics* 12, 339–346.

Koop, G. and M. F. J. Steel. (2001). Bayesian Analysis of Stochastic Frontier Models. Chapter 24 in B. Baltagi (ed.), *A Companion to Theoretical Econometrics*. Oxford: Blackwell, 520–537.

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

Meeusen, W. and J. van den Broeck. (1977). "Efficiency Estimation from Cobb-Douglas Productions with Composed Errors." *International Economic Review* 8, 435–444.

Ritter, C. and L. Simar. (1997). "Pitfalls of Normal-Gamma Stochastic Frontiers and Panel Data." *Journal of Productivity Analysis* 8, 167–182.

Tsionas, E. G. (2002). "Stochastic Frontier Models with Random Coefficients." *Journal of Applied Econometrics* 17, 127–147.

Tsionas, E. G. (2000). "Full Likelihood Inference in Normal-Gamma Stochastic Frontier Models." *Journal of Productivity Analysis* 13, 183–205.

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.