Comparing Bayesian Models for Production Efficiency

Ricardo S. Ehlers*

Universidade Federal do Paraná, Brazil

April, 2007

Abstract

In this paper we use Markov chain Monte Carlo (MCMC) methods in order to estimate and compare stochastic production frontier models from a Bayesian perspective. We consider a number of competing models in terms of different production functions and the distribution of the asymetric error term. All MCMC simulations are done using the package `JAGS` (Just Another Gibbs Sampler), a clone of the classic `BUGS` package which works closely with the `R` package where all the statistical computations and graphics are done.

Key Words: Markov chain Monte Carlo, Gibbs sampler, JAGS, model comparison.

*Address for correspondence:Departamento de Estatística, Universidade Federal do Paraná, Caixa Postal 19081. 81531-990, Curitiba - PR, Brazil. E-mail: ehlers@leg.ufpr.br

1 Introduction

In stochastic production frontier models it is usually assumed that the error term is composed of a random error (v) capturing statistical noise and a one-sided non-negative error (u). For N firms observed, the model can be expressed as

$$y_i = f(\boldsymbol{x}_i, \boldsymbol{\beta}) + v_i - u_i,\ i = 1, \ldots, N$$

where y_i is the logarithm of an output, $\boldsymbol{x}_i$ is a vector of the logarithms of inputs including an intercept and possibly crossproducts, $\boldsymbol{\beta}$ is the vector of coefficients and v_i are independent and identically distributed $N(0, \sigma_v^2)$ error terms and assumed to be independent of the non-negative random variable u_i. Thus, the total error term $\epsilon_i = v_i - u_i$ has an asymmetric distribution which can be obtained as

$$p(\epsilon) = \int_0^\infty p(u) p_N(\epsilon + u | 0, \sigma_v^2) du.$$

where $p_N(\cdot|\mu, \sigma^2)$ stands for the $N(\mu, \sigma^2)$ density function. Technical efficiency of the i-th firm is measured by $r_i = \exp(-u_i)$ and u_i is assigned a non-negative asymmetric distribution so that $r_i \in (0, 1)$.

In practice, there is also a great deal of uncertainty with respect to both the production function $f(\cdot)$ and the distribution of u_i. In particular, a number of choices for the distribution of u_i have been proposed in the literature. For example, in their pioneering work in this area Meeusen and van der Broeck (1977) and Aigner, Lowell, and Schmidt (1977) used exponential and half-Normal distributions respectively. Later proposals include a $N(\xi, \lambda)$ truncated to $u_i > 0$ allowing both mean and variance to be estimated (Stevenson 1980), Gamma distributions (Greene 1990, Koop, Steel, and Osiewalski 1995) and Log-Normal distributions (Migon and Medrano 2004).

Each of these choices can induce very different behaviours in the distribution of the technical efficiencies r_i and more recently Griffin and Steel (2004) proposed a Generalized Gamma distribution for u_i. In this case, its density function is given by

$$p(u_i | c, \phi, \lambda) = \frac{c\lambda^\phi}{\Gamma(\phi)} u_i^{c\phi - 1} \exp(-\lambda u_i^c),\ \phi > 0,\ \lambda > 0 \tag{1}$$

and it is easy to see that most of the aforementioned proposals are simpler cases of this three-parameter family (except the truncated and half-Normal). So, we have a Gamma distribution when $c = 1$, an Exponential distribution when $c = \phi = 1$, and a Log-Normal distribution as a limiting case. Also, a Weibull distribution is obtained when $\phi = 1$.

Production Functions

Suppose that we want to relate values of a product (output) Q produced with 2 factors (inputs): capital (K) and Labour (L). Let $y = \log(Q)$, $x_1 = \log(K)$ and $x_2 = \log(L)$. Two commonly used production functions are, the Cobb-Douglas which is linear in the logarithms of the inputs, i.e.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2,$$

where $\beta_0 \in \mathbb{R}$, $\beta_1 > 0$ and $\beta_2 > 0$, and the Translog which includes squares and cross-products,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \frac{1}{2}\beta_3 x_1^2 + \frac{1}{2}\beta_4 x_2^2 + \beta_5 x_1 x_2$$

where $\beta_0 \in \mathbb{R}$ and $\beta_i > 0$, $i = 1, \ldots, 5$. So, of course the Cobb-Douglas is a special case of the Translog function for $\beta_3 = \beta_4 = \beta_5 = 0$. More recently, Migon and Medrano (2004) investigated two other types of production function namely the Constant Elasticity of Substitution (CES) which is non-linear in the inputs

$$y = \beta_0 + v \log(\, [(1-\delta)K^{-\rho} + \delta L^{-\rho}]^{-1/\rho})$$

where $0 < \delta < 1$, $-\infty < v < \infty$ and $\rho > -1$, and the Generalized Production Function (GPF), first proposed by Zellner and Revankar (1970), where

$$z = y + \lambda \exp(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2$$

where $\beta_0 \in \mathbb{R}$, $\lambda > 0$, $\beta_1 > 0$ and $\beta_2 > 0$.

In this paper, we focus on the comparison of these production functions and distributions of the efficiency term. We adopt a Bayesian approach and use Markov

Chain Monte Carlo (MCMC) simulation to estimate parameters and compare models. The MCMC sampling is done using the package `JAGS` (Just Another Gibbs Sampler, Plummer 2003) which was originally developed as a clone of the classic `BUGS` package. JAGS was written in `C++` and is designed to work closely with the `R` package where all statistical computations and graphics are done. This package is open source and freely available and can be downloaded from the website

`http://www-fis.iarc.fr/~martyn/software/jags/`.

A compiled version for Windows users is also available from this website.

We provide a set of `R` functions to build the required files and analyse the output using the functionality from the `R` package `coda`. So, the user can choose the production function and the distribution of the efficiency term, as well as provide data, prior parameters and starting values for the MCMC chains. The main idea is that `JAGS` can be used as a sampling device in `R` and the user does not need to know the `BUGS` syntax (as when using the `WinBUGS` package). However, the model specification, prior distributions and initial values are written in external files so that they can indeed be modified by the user. Initial values which are not explicitly set by the user are automatically set by `JAGS`. All `R` functions used in this paper are available from the author at `http://leg.ufpr.br/~ehlers/SPF`

The rest of the paper is organized as follows. In Section 2 the model classes which can be fitted are described in more details. In Section 3 a short overview on MCMC-based model comparison issues is given. Empirical results are provided in Section 4 and we give a short description of the implementation using `R`. We close the paper in Section 5 with a general discussion on possible extensions of the models adopted and some possibilities for future work.

2 Bayesian Models

Stochastic frontier models have been analysed from a Bayesian viewpoint by many authors recently, especially using MCMC methods for estimation. See for example

Koop, Steel, and Osiewalski (1995), Fernandez, Osiewalski, and J. (1997), Osiewalski and Steel (1998) and (Migon and Medrano 2004).

For the Cobb-Douglas, Translog and CES production functions the likelihood function is given by

$$p(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} p_N(y_i|f(x_{i1}, x_{i2}, \boldsymbol{\theta}) - u_i, \sigma^2)$$

where $\boldsymbol{\theta}$ contains the parameters defining the production functions and $p_N(y|\mu, \sigma^2)$ denotes the normal probability density function with mean μ and variance σ^2. However, for the generalized production function we need to compute the Jacobian of the transformation from Z to Y, so the likelihood function is now given by

$$p(\boldsymbol{y}|\boldsymbol{\theta}) = \prod_{i=1}^{n} p_N(z_i|f(x_{i1}, x_{i2}, \boldsymbol{\theta}) - u_i, \sigma^2) \left|\frac{dz_i}{dy_i}\right|. \tag{2}$$

We also need to specify prior distributions for the parameters in $\boldsymbol{\theta}$ and of course this depends on which production function is used. Whatever the production function these parameters are all assumed to be *a priori* independent. The intercept is assigned an unrestricted normal prior around zero, i.e. $\beta_0 \sim N(0, \sigma^2_\beta)$ while the other coefficients are constrained to be positive as we want to exclude production frontiers in which more inputs lead to less output. In this paper we assign priors $\beta_j \sim N(0, \sigma^2_\beta)$ truncated to $\beta_j > 0$. Finally, for a CES production function we assign priors $v \sim N(0, \sigma^2_v)$, $\rho \sim N(0, \sigma^2_\rho)$ truncated to $\rho > -1$ and $\delta \sim Be(a, b)$ where $Be(a, b)$ stands for a Beta distribution. Also, $\sigma^{-2} \sim Ga(a, b)$ where $Ga(a, b)$ stands for a Gamma distribution with mean a/b and variance a/b^2.

In this paper, a number of choices for the distribution of the inefficiency terms is also to be compared. These are listed below, together with the associated hyperpriors.

1. Assuming that $u_i \sim Exp(\lambda)$ then assigning a prior $\lambda \sim Exp(-\log r^*)$ implies that the prior median efficiency equals r^*. van den Broeck et al. (1994) show that this leads to a proper but relatively non-informative prior on the firm efficiencies.

2. Assuming a half-normal distribution, i.e. $u_i \sim N(0, \lambda)$ truncated to $u_i > 0$ then we assign a prior $\lambda^{-1} \sim Ga(1, 1/37.5)$. van den Broeck et al. (1994) show that this leads to a prior median efficiency of approximately 0.875 with a reasonable spread.

3. When a truncated normal distribution is assumed, i.e. $u_i \sim N(\xi, \lambda)$ truncated to $u_i > 0$ and we include a prior $\xi \sim N(0, \sigma_\xi^2)$ and use the same prior for λ as above.

4. Assuming that $u_i \sim Ga(\phi, \lambda)$, Griffin and Steel (2004) propose a prior on ϕ and λ which extends the informative prior for an exponential inefficiency distribution. Assigning $\phi^{-1} \sim Ga(d_1, d_1 + 1)$ implies that ϕ has a prior mode at one. So this prior is centred around an exponential prior with d_1 controlling the variability of ϕ. Griffin and Steel (2004) suggest taking $d_1 = 3$ and this is the value adopted here. We also take $\lambda|\phi \sim Ga(\phi, -\log r^*)$ where r^* is the prior median inefficiency.

5. When a log-normal distribution is assigned for the u_i we consider a sub-family with location parameter equal to zero and the scale parameter can be chosen to make the prior relatively vague. So, $u_i \sim LN(0, \psi^2)$ and we assign a hyperprior $\psi^{-2} \sim Ga(a, b)$.

6. Assigning a Generalized Gamma distribution for the u_i the parameters c, ϕ and λ in (1) are specified here following the development in Griffin and Steel (2004). We take $\lambda|c, \phi \sim Ga(\phi, (-\log r^*)^c)$ and assume prior independence between c and $\psi = \phi c$ assigning priors

$$\psi^{-1} \sim Ga(d_1, d_1 + 1) \quad \text{and} \quad c^{-1} \sim Ga(d_2, d_2 + 1).$$

   This prior is again centred over the exponential case (mode at $\phi = c = 1$) but allows considerable deviations from the exponential if we choose d_1 and d_2 not too large. We take $d_1 = d_2 = 3$ here.

7. Finally, assuming that $u_i \sim Weibull(c, \lambda)$ we adapt the development in item (6) with $\phi = 1$ and assign $\lambda|c \sim Exp((-\log r^*)^c)$ and $c^{-1} \sim Ga(d_2, d_2 + 1)$ which is centred over the exponential case. We take $d_2 = 3$ here.

The `R` code implemented allows the user to specify values for the hyperparameters σ^2_β, σ^2_v, σ^2_ρ, a, b, d_1, d_2 and r^*.

Ranking the Firms

Samples of the firm-specific efficiency terms (e^{-u_i}) are immediately generated by the sampler. In practice, the investigator is often more interested in the ranks of these terms and their posterior distributions are also readily available from the MCMC output. In the jth iteration, values of $\exp(-u_i^{(j)})$, $i = 1, \ldots, N$ are generated and by ordering these values then the position of the ith value in the ordered set is the rank of the ith firm (a high rank corresponds to high efficiency). After convergence we have a sample of ranks for each firm which can be summarized to provide an estimate of the mean or median rank for each firm plus a 100(1-α)% credible interval which captures the uncertainty associated with the rank position of each firm. Computationally, it is faster to obtain this last sample using `R` commands on the sampled values of $\exp(-u_i)$.

3 Comparing Models

In order to compare and select the most appropriate model among those considered here we use the *Deviance Information Criterion* (DIC), where lower values indicate a good model fit relative to the number of parameters in the model (for full details see Spiegelhalter et al. 2002). Denoting the competing models by $M_1, M_2, \ldots$ and the vector of parameters under model M_i by ξ_i the criterion is defined as $DIC(M_i) = \overline{D}_i + p_i$ where $D_i = -2\log(p(\boldsymbol{y} \mid \xi_i, M_i))$ is the deviance, measuring model fit, and $\overline{D}_i$ is the posterior mean of D_i. The penalizing term p_i measures model complexity and is given by $\overline{D}_i - D(\bar{\xi}_i)$ where $\bar{\xi}_i$ is the posterior mean of ξ_i.

Routine application of DIC for model determination has become standard practice in Bayesian applied work as it is easily computed during the simulation of the Markov chains. In this work we include an `R` function for the computation of DIC and the user can compare models on the basis of this criterion. However, it could be misleading just to report the model with the lowest DIC as it is difficult to say what would constitute an important difference between DIC values. Also, the DIC is subject to Monte Carlo sampling error since it is a function of stochastically simulated quantities. This might cast some doubt whether an improvement in model fit is substantial. One way around this problem, is to use DIC weights obtained by subtracting from each DIC the value associated with the "best" model and then setting

$$w_i \propto \exp(-\Delta DIC(M_i)/2)$$

where $\Delta DIC(M_i)$ denotes the transformed DIC value for model i. The weights are then normalized to sum to 1 over the models under consideration. This approach was first suggested in Burnham and Anderson (1998) Section 4.2 for the Akaike's Information Criterion (AIC) where the differences are interpreted as the strength of evidence. The approach can be extended to be used with the DIC which, as pointed out in Spiegelhalter et al. (2002), can be viewed as a Bayesian analogue of AIC.

4 Empirical Results

In this section we analyse 123 cross-sectional data from the US electric industry in 1970. The data was originally analysed by Christensen and Greene (1976) and subsequently by Greene (1990) and Griffin and Steel (2007). Here, the production function relates the output produced to labour and capital corrected by their respective factor prices. The data can be obtained at

`http://econ.queensu.ca/jae/1998-v13.2/zellner-ryu/data.zr`.

For each model a total of 100 000 iterations were run, and the first half was discarded as burn-in. This was intended to be a conservative burn-in as the diagnostics of Raftery and Lewis (1992) indicated that convergence has actually occurred earlier. After the burn-in period the simulated values of every 5th iteration were kept for posterior analysis. So, parameter estimates and model comparison were based on approximate samples of size 10 000 from the posterior distributions. The diagnostics of Geweke (1992) and an analysis of the behavior of the chains along the iterations did not indicate lack of convergence. From this sample several characteristics of the posterior distribution may be estimated, and the main interest here is to provide estimates of the mean and median rank for the firms plus 100(1-α)% credible intervals.($d1 = d2 = 3$, $r^* = 0.80$)

In Table 1 the computed values of $\overline{D}_i$, $D(\bar{\xi}_i)$, the penalizing terms p_i the DIC and the DIC weights for each model are presented. The last column shows their associated rankings. We can see that the most adequate model is the one that uses the generalized production function and a truncated normal distribution for the inefficiency terms. We note however that model comparison is much easier in the transformed scale of the DIC weights. We can see that the weight for the second best model (Constant Elasticity of Substitution + truncated normal inefficiencies) is less than half the weight for the most adequate model. Another subset of models (GPF + lognormal, CES + lognormal and Cobb-Douglas + truncated normal) received small weights 0.07121, 0.06074 and 0.05449. All other models receive very small weights. It is also worth noting that, within each production function, the truncated normal distribution for the inefficiencies is always selected.

A feature of interest is the ability of each model to classify the firms, in particular to differenciate between the more efficient firms from the less efficient ones. In terms of MCMC simulations this is accomplished by comparing the posterior distributions of the efficiency terms for the best and worst firms. Figure 1 shows the posterior distributions of the ranks associated with the efficiency measures of the best firm (left column) and the worst firm (right column) using the truncated normal distribution considering the four production functions. As we can see, the truncated normal frontier model differentiates pretty well the more efficient from the less efficient firms.

Table 1 about here

Figure 1 about here

In Appendix A we show the JAGS commands used to estimate the selected model (generalized production function and a truncated normal distribution for the inefficiency terms). This is for illustration purposes only since the user does not need to know the JAGS commands syntax. We note that we center the regressors around their sample averages in order to improve convergence. The intercept is set back to its original scale in the end. Also, we have to use the so called zeros trick in which artificial data with all values equal to zero are assigned a Poisson distribution with mean given by $-\log(p(\boldsymbol{y}|\theta))$ where $p(\boldsymbol{y}|\theta)$ is given in (2). This trick is also commonly used by WinBUGS users.

5 Discussion

In this work we adopted a Bayesian approach to estimate and compare stochastic production frontier models The method was illustrated with a real data example. Estimates of the posterior distribution were obtained via MCMC methods where inference is based on an approximate sample from the posterior distribution.

A set of production functions and distributions for the inefficiency terms which might be of potential influence on the ranking of the firms were compared via DIC. We provide the R code which builds the required files to run the JAGS package and to analyse the output using the functionality from the R package coda. It is worth noting that the user is not required to know the BUGS syntax to run these R functions. The code was implemented for cross-sectional data but can be easily extended for analysing more complicated models, e.g. balanced and unbalanced panel data, inclusion of covariates in the inefficiency distributions and models with time-varying inefficiencies. Also, the symmetric term (v_i) could be assigned a a Student-t distribution with unknown degrees of freedom in other to robustify the model. Care should be exercised when specifying a prior distribution for the number of degrees of freedom (Migon, Rosa, and Fonseca 2005 derived a reference prior for this parameter).

Finally, an area still to be explored in more depth in the literature is model comparison and selection via trans-dimensional MCMC algorithms (e.g. reversible jump MCMC). Surely the only sensible approach is to calculate posterior model probabilities and there is still considerable room for more research in this area. All these extensions are currently being investigated by the author.

References

Aigner, D., C. Lowell, and P. Schmidt (1977). Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics 6*, 21–37.

Burnham, K. P. and D. R. Anderson (1998). *Model Selection and Inference: A Practical Information-Theoretic Approach*. Springer: New York.

Christensen, L. R. and W. H. Greene (1976). Economies of scale in U.S. electric power generation. *Journal of Political Economy 84*(4), 655–681.

Fernandez, C., J. Osiewalski, and S. M. F. J. (1997). On the use of panel data in stochastic frontier models with improper priors. *Journal of Econometrics 79*, 169–193.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, A. F. M. Smith, A. P. Dawid, and J. O. Berger (Eds.), *Bayesian Statistics 4*, pp. 169–193. New York: Oxford University Press.

Greene, W. H. (1990). A gamma-distributed stochastic frontier model. *Journal of Econometrics 46*, 141–163.

Griffin, J. E. and M. F. J. Steel (2004). Flexible mixture modelling of stochastic frontiers. Technical Report 418, University of Warwick.

Griffin, J. E. and M. F. J. Steel (2007). Bayesian stochastic frontier analysis using WinBUGS. *Journal of Productivity Analysis 27*. To appear.

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 der Broeck (1977). Efficiency estimation from Cobb-Doublas production functions with composed errors. *International Economic Review 8*, 435–444.

Migon, H. S. and L. A. T. Medrano (2004). Deviance-based criteria for comparing Bayesian stochastic production frontier models (in portuguese). Technical Report 176, Universidade Federal do Rio de Janeiro.

Migon, H. S., M. A. F. Rosa, and T. C. O. Fonseca (2005). Bayesian reference analysis for the class of generalized hyperbolic distributions. In *Latin American Congress on Bayesian Statistics - Cobal II.*

Osiewalski, J. and M. F. J. Steel (1998). Numerical tools for the Bayesian analysis of stochastic frontier models. *Journal of Productivity Analysis 10*, 103–117.

Plummer, M. (2003). JAGS: A program for analysis Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, and A. Zeileis (Eds.), *Proceedings of the 3rd International Workshop on Distributed Statistical Computing.* Technische Universitaet Wien, Vienna, Austria.

Raftery, A. E. and S. M. Lewis (1992). One long run with diagnostics: implementation strategies for Markov chain Monte Carlo. discussion of Gelman and Rubin (1992). *Statistical Science 7*, 493–497.

Spiegelhalter, D. J., N. G. Best, B. P. Carlin, and A. van der Linde (2002). Bayesian measures of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B 64*, 1–34.

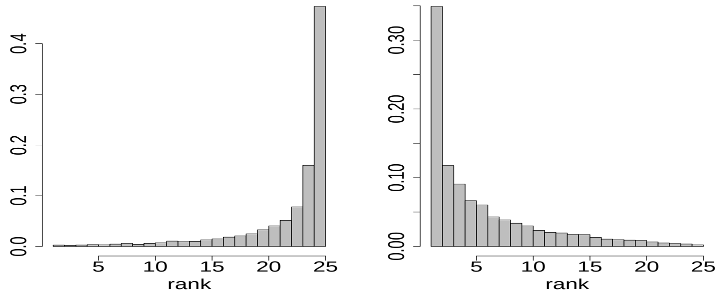
Stevenson, R. E. (1980). Likelihood functions for generalized stochastic frontier estimation. *Journal of Econometrics 13*, 57–66.

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.

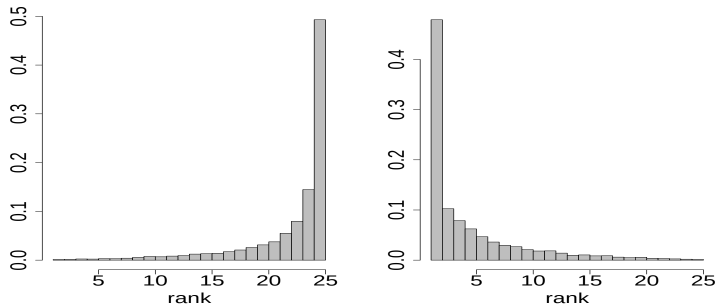
Zellner, A. and N. S. Revankar (1970). Generalized production functions. *The Review of Economic Studies 36*, 241–250.

Table 1: $\overline{D}_i$, $D(\bar{\xi}_i)$, penalizing terms (p_i), DIC values, computed DIC normalised weights and rank for each model obtained based on 10 000 simulations.

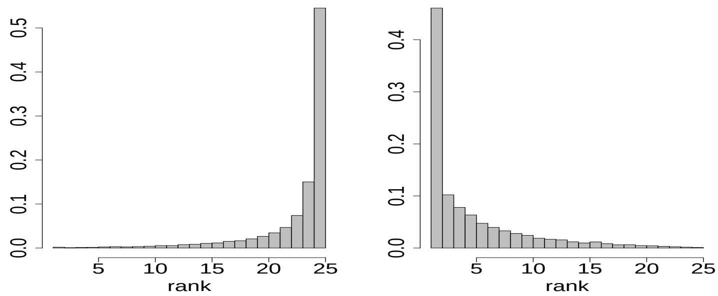
Prod Function	inefficiency	$\overline{D}_i$	$D(\bar{\xi}_i)$	p_i	DIC	weight	rank
Cobb-Douglas	exp	−19.5609	−28.3636	8.8027	−10.7581	0.00044	23
Cobb-Douglas	tnorm	−30.9996	−41.6037	10.6041	−20.3956	0.05449	5
Cobb-Douglas	halfnorm	−19.0528	−27.2502	8.1974	−10.8554	0.00046	22
Cobb-Douglas	gamma	−19.8941	−28.6195	8.7254	−11.1687	0.00054	21
Cobb-Douglas	gen.gamma	−20.2863	−28.6803	8.3940	−11.8924	0.00078	19
Cobb-Douglas	weibull	−18.6541	−26.1266	7.4726	−11.1815	0.00054	20
Cobb-Douglas	lognorm	−27.6661	−37.6017	9.9355	−17.7306	0.01437	6
Translog	exp	−15.0361	−23.8715	8.8353	−6.2008	0.00005	25
Translog	tnorm	−23.9300	−33.2167	9.2867	−14.6432	0.00307	7
Translog	halfnorm	−12.8873	−20.6193	7.7320	−5.1554	0.00003	27
Translog	gamma	−15.1713	−23.7525	8.5812	−6.5900	0.00005	24
Translog	gen.gamma	−13.8555	−21.7858	7.9303	−5.9252	0.00004	26
Translog	weibull	−12.8305	−20.5585	7.7280	−5.1025	0.00003	28
Translog	lognorm	−22.0603	−31.4122	9.3519	−12.7084	0.00117	13
CES	exp	−19.5195	−26.4384	6.9190	−12.6005	0.00111	15
CES	tnorm	−31.2346	−39.2628	8.0282	−23.2065	0.22216	2
CES	halfnorm	−19.5870	−25.8898	6.3028	−13.2842	0.00156	11
CES	gamma	−20.2690	−26.6339	6.3649	−13.9041	0.00212	8
CES	gen.gamma	−20.1451	−26.8770	6.7319	−13.4132	0.00166	10
CES	weibull	−18.8055	−24.9562	6.1507	−12.6548	0.00114	14
CES	lognorm	−28.1078	−35.6027	7.4949	−20.6130	0.06074	4
GPF	exp	−19.9535	−27.9918	8.0382	−11.9153	0.00078	18
GPF	tnorm	−33.4537	−41.8647	8.4110	−25.0428	0.55643	1
GPF	halfnorm	−20.0239	−27.4804	7.4566	−12.5673	0.00109	16
GPF	gamma	−20.9083	−28.9271	8.0188	−12.8895	0.00128	12
GPF	gen.gamma	−20.8332	−28.1687	7.3355	−13.4978	0.00173	9
GPF	weibull	−19.6832	−27.0591	7.3759	−12.3073	0.00095	17
GPF	lognorm	−28.8473	−36.7637	7.9164	−20.9309	0.07121	3



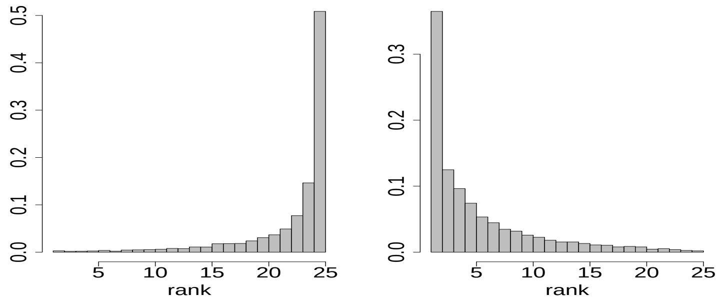
(a) Cobb-Douglas



(b) Translog



(c) CES



(d) GPF

Figure 1: Posterior distributions of the ranks of the best firm (left column) and the worst firm (right column) for a truncated normal inefficiency term considering the four production functions.

A JAGS commands

JAGS commands used for the model with a generalized production function and a truncated normal distribution for the inefficiency terms.

```
data {
  for (i in 1:N) {
    zeros[i] <- 0
    X[i,1] <- Xreg[i,1]-xbar[1]
    X[i,2] <- Xreg[i,2]-xbar[2]
  }
  xbar[1] <- mean(Xreg[,1])
  xbar[2] <- mean(Xreg[,2])
}
model {
  for(i in 1:N) {
    zeros[i] ~ dpois(p[i])
    p[i] <- -0.5*log(tau/(2*3.141593))
    +0.5*tau*pow(y[i]+gamma*exp(y[i])-mu[i],2)
    -log(1+gamma*exp(y[i]))+10000
    mu[i] <- alpha0 +inprod(beta[],X[i,]) - u[i]
  }
  beta[1] ~ dnorm(0.0,0.001)T(0,)
  beta[2] ~ dnorm(0.0,0.001)T(0,)
  alpha0 ~ dnorm(0.0,0.001)
  gamma ~ dgamma(0.1,0.01)
  tau ~ dgamma(0.01,0.01)
  sigma2 <- 1/tau
  for (i in 1:N) {
    u[i] ~ dnorm(zeta,invlambda)T(0,)
    eff[i] <- exp(-u[i])
  }
  invlambda ~ dgamma(1,1/37.5)
  zeta ~ dnorm(0.0, 1E-3)T(0,)
  lambda <- 1/invlambda
  alpha <- alpha0-inprod(beta[],xbar[])}
```