Bayesian Quantile Stochastic Frontier Models

Angel Arroyo Hinostroza, Ralph dos Santos Silva, Helio dos Santos Migon
Instituto de Matemática, Universidade Federal de Rio de Janeiro, Brazil

Abstract

In this paper we propose a new approach to explore the stochastic frontier models which uses the power of Bayesian quantile regression. Compared with usual models based on regression in the conditional mean, our proposal inherits the advantages of quantile regression, such as robustness of estimators as it does not need to assume any distribution to the data nor assume homoscedasticity. Moreover, it also brings more details to the analyst since that several quantiles provide more information about the stochastic frontier. In addition, our proposal allows a better comparison of technical efficiency estimation among firms through analysis of several quantiles in the stochastic frontier.
Keywords: Bayesian quantile regression, Gibbs sampling, Stochastic Frontier, Technical efficiency.

1. Introduction

In economic theory, aspects of productive systems based in production functions may be described by *Stochastic Frontier* models, introduced initially by Aigner et al. (1977) and Meeusen and van den Broeck (1977). These models have been extensively applied in the study of technical efficiencies of firms. Many proposals in the literature have been established in what might be considered as a conditional mean environment, exploring solely the relationship between the inputs and the conditional mean of output.

On the other hand, in recent years several works have explored the quantile estimation in stochastic frontier models from a classical approach. Like Bernini et al. (2004) that fo-

Email addresses: angel@dme.ufrj.br (Angel Arroyo Hinostroza), ralph@im.ufrj.br (Ralph dos Santos Silva), migon@im.ufrj.br (Helio dos Santos Migon)

cused their attention on the composed error quantile estimation, and also the estimation of technical efficiency. Or Liu et al. (2008) that through a Monte Carlo simulation suggested that compared to non-parametric *data envelopment analysis* (DEA) and parametric *stochastic frontier analysis* (SFA), the quantile regression approach is better (depending on the quantile estimated) in terms of their ability to accurately estimate efficiencies. Behr (2010) discussed the use of quantile regression estimation as a robust alternative to estimate the production frontier, considered as a quantile close to one of the composed error, and show that even when generating data according to the assumptions of the stochastic frontier model, efficiency estimates obtained from quantile regressions resemble to efficiency estimates obtained from SFA. Moreover, Jradi et al. (2019) affirmed that the stochastic frontier corresponds explicitly to a specific quantile of the output distribution and provide a computational approach to recover this quantile. In their work the estimation of technical efficiency is again ignored.

We propose, from a Bayesian approach, a new form to apply the quantile estimation in the stochastic frontier models. Instead of exploring the relationship between inputs and output quantile, we explore the relationship between inputs and quantiles of efficient output (i.e., the output without inefficiencies). Our proposal inherits the whole power of quantile regression, which, in contrast to the conditional mean environment, provides more robust estimators, does not need to assume homoscedasticity - this feature allows for increasing the flexibility of the model - and no need to specify the likelihood (a distribution for the error term). See Li (2015) for a discussion about the advantages of quantile regression models.

The outline of this paper is organized as follows. A short background about Bayesian quantile regression models and stochastic frontier models is given in Section 2. Our proposed model and Bayesian inference is shown in Section 3. Monte Carlo studies are reported in Section 4, and real data applications are presented in Section 5. Section 6 concludes.

2. Review

2.1. Bayesian quantile regression model

Let y_i be the continuous response variable and $\mathbf{x}_i$ be the p-dimensional vector of covariates for the i-th observation ($i = 1, \ldots, \mathrm{N}$). For some $\tau \in (0,1)$, $\mathcal{Q}_\tau(y_i \mid \mathbf{x}_i)$ denote the τ-th quantile regression function of y_i given $\mathbf{x}_i$. Suppose that there exists a linear relationship between $\mathcal{Q}_\tau(y_i \mid \mathbf{x}_i)$ and $\mathbf{x}_i$, i.e., $\mathcal{Q}_\tau(y_i \mid \mathbf{x}_i) = \mathbf{x}_i^\top \boldsymbol{\beta}_\tau$, where $\boldsymbol{\beta}_\tau$ is a p-dimensional parameter vector.

The quantile regression model is usually expressed by

$$y_i = \mathbf{x}_i^\top \boldsymbol{\beta}_\tau + \varepsilon_i, \quad i = 1, \ldots, \mathrm{N}, \tag{1}$$

where ε_i's are independent and identically distributed (i.i.d.) error terms restricted to have the τ-quantile equal to zero, i.e., $\Pr(\varepsilon_i \leqslant 0) = \tau$. The error distribution is often left unspecified and the estimation for $\boldsymbol{\beta}$ (the τ has been dropped from $\boldsymbol{\beta}_\tau$ to simplify the notation) proceeds by minimizing

$$S(\boldsymbol{\beta}) = \textstyle\sum_i \rho_\tau(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}), \tag{2}$$

where $\rho_\tau(\varepsilon) = \varepsilon\{\tau - \mathbb{I}_{(-\infty,0)}(\varepsilon)\}$ is called the *quantile loss (or check) function*, and $\mathbb{I}_A(.)$ denotes the usual indicator function of the subset A. Details about quantile regression from a classical approach can be found in Koenker (2005).

From a Bayesian perspective, Yu and Moyeed (2001) proposed to model the error term, ε_i, in Equation (1) with the asymmetric Laplace distribution, denoted by $\mathcal{LA}_\tau(\mu, \sigma)$, where $\tau \in (0,1)$, $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$ are the quantile, the location and the scale parameters, respectively. In their proposal, they kept τ known and fixed. Moreover, based on a random sample, their construction led to the (auxiliary) likelihood function given by

$$L(\boldsymbol{\beta}, \sigma) = \frac{\tau^{\mathrm{N}}(1-\tau)^{\mathrm{N}}}{\sigma^{\mathrm{N}}} \exp\left\{ -\frac{1}{\sigma} \textstyle\sum_i \rho_\tau(y_i - \mathbf{x}_i^\top \boldsymbol{\beta}) \right\}. \tag{3}$$

Considering improper uniform priors for $\boldsymbol{\beta}$ (and $\sigma = 1$), Yu and Moyeed (2001) proved that the posterior distribution of $\boldsymbol{\beta}$ is proper although it has no closed form. They employed the

Metropolis algorithm, with slow convergence, to draw a sample from the posterior distribution. Additionally, Sriram et al. (2013) explored the consistency of the quantile regression estimator based on the auxiliary likelihood in Equation (3), and proposed an approximation to the posterior distribution for large samples, specially to reduce the bias on the posterior covariance matrix.

Using a location-scale mixture of normals representation of the asymmetric Laplace distribution (with $\sigma = 1$) and a Gaussian prior for $\boldsymbol{\beta}$, Kozumi and Kobayashi (2011) showed that it is possible to get full conditional distributions in closed form which are, in turn, easy to sample from. This strategy allows the implementation of a computationally efficient Gibbs sampler algorithm. It can also be easily extended to include σ estimation.

A random variable ε with $\mathcal{LA}_\tau(0, \sigma)$ distribution admits a location-scale mixture representation $\kappa_1 \nu + z\sqrt{\kappa_2 \sigma \nu}$, where $\kappa_1 = (1-2\tau)/[(\tau(1-\tau)]$ and $\kappa_2 = 2/[\tau(1-\tau)]$ are constants, ν is a latent variable with exponential distribution with mean σ, denoted by $\nu \sim \mathcal{E}(\sigma)$, and z is a random variable with standard normal distribution, denoted by $z \sim \mathcal{N}(0,1)$. More information about properties of asymmetric Laplace distribution and its representations can be found in Kotz et al. (2001).

Then, using the result above and Equation (1), the Bayesian quantile regression model can alternatively be constructed by

$$\begin{aligned}
(y_i \mid \boldsymbol{\beta}, \sigma, \nu_i) &\sim \mathcal{N}(\mathbf{x}_i^\top \boldsymbol{\beta} + \kappa_1 \nu_i, \kappa_2 \sigma \nu_i) \\
(\nu_i \mid \sigma) &\sim \mathcal{E}(\sigma), \quad i = 1, \ldots, \mathrm{N}.
\end{aligned}$$

The Bayesian model specification is completed by choosing the prior distribution as $\pi(\boldsymbol{\beta}, \sigma) = \pi(\boldsymbol{\beta})\pi(\sigma)$, with $\boldsymbol{\beta}$ having a p-variate Gaussian distribution with mean $\mathbf{m}_0$ and covariance matrix $\mathbf{C}_0$, denoted by $\boldsymbol{\beta} \sim \mathcal{N}_p(\mathbf{m}_0, \mathbf{C}_0)$, and σ as an inverse gamma distribution, denoted by $\sigma \sim \mathcal{IG}(n_0, s_0)$, where n_0 and s_0 represent its shape and scale parameters, respectively.

The posterior distribution of $(\boldsymbol{\beta}, \sigma)$ is obtained under a data augmentation scheme

$\pi(\boldsymbol{\beta}, \sigma \mid \mathbf{y}) \propto \int \pi(\boldsymbol{\beta}, \sigma, \boldsymbol{\nu} \mid \mathbf{y}) d\boldsymbol{\nu}$ where

$$\pi(\boldsymbol{\beta}, \sigma, \boldsymbol{\nu} \mid \mathbf{y}) \propto \left[\prod_i f(y_i \mid \boldsymbol{\beta}, \sigma, \nu_i)\right] \left[\prod_i \pi(\nu_i \mid \sigma)\right] \pi(\boldsymbol{\beta})\pi(\sigma), \tag{4}$$

with $\mathbf{y} = (y_1, \ldots, y_{\mathrm{N}})^\top$ and $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_{\mathrm{N}})^\top$. It is straightforward to prove that the resulting full conditional distributions from Equation (4) are Gaussian for $(\boldsymbol{\beta} \mid \mathbf{y}, \boldsymbol{\nu}, \sigma)$, inverse gamma for $(\sigma \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\beta})$ and generalized inverse Gaussian (Jørgensen, 1982) for $(\nu_i \mid y_i, \boldsymbol{\beta}, \sigma)$, for $i = 1, \ldots, \mathrm{N}$. This allows for the implementation of the Gibbs sampler algorithm to draw a sample from the posterior distribution.

However, the consistency showed by Sriram et al. (2013) does not guarantee adequate credibility interval estimates for the quantile regression parameters. The employment of the auxiliary likelihood in Equation (3), based on the asymmetric Laplace distribution, generates underestimation of the posterior variances of the quantile regression parameters when compared to a possibly true error distribution different from the asymmetric Laplace. This feature is exacerbated when considering extremes quantiles (away from the median). Yang et al. (2016) described this problem and developed, under conditions that guarantee the asymptotic normality of the quantile regression estimators, an adjustment in the posterior covariance matrix to obtain asymptotically-based interval estimates. We pursue this idea later on.

2.2. Stochastic frontier models

To define the elements of stochastic frontier models more generally (panel data model) let y_{it} be the output quantity and $\mathbf{x}_{it}$ be the p-dimensional vector of inputs quantities for the i-th firm ($i = 1, \ldots, \mathrm{N}$) at time t ($t = 1, \ldots, \mathrm{T}$). The case $\mathrm{T} = 1$ corresponds to cross-sectional model, in which case the input and output quantities are denoted by y_i and $\mathbf{x}_i$, respectively. Let $f(\mathbf{x}_{it}; \boldsymbol{\beta})$ be the functional form which relates the output and input quantities, where $\boldsymbol{\beta}$ is a parameter vector. In particular, the production function can be considered as functional form for simplicity. Many common production functions are listed in Coelli et al. (2005, Table 8.1).

The seminal stochastic frontier model, proposed by Aigner et al. (1977) and Meeusen and van den Broeck (1977), is given by

$$y_i = f(\mathbf{x}_i; \boldsymbol{\beta}) + \varepsilon_i - \omega_i, \qquad i = 1, \ldots, \mathrm{N}. \tag{5}$$

In this modeling the production process is subject to two random disturbances. The first one, denoted by ω_i, is defined as non-negative disturbance which generates a decrease in the firms output. This disturbance is called *technical inefficiency* and it is the result of factors under the firms control. The second disturbance, denoted by ε_i, is defined as a random error term without signal constraint which measure the effect of uncontrolled variables, and often considered with Gaussian distribution, although there are proposals considering robustness through the Student-t distribution (Griffin and Steel, 2007).

There is a vast literature discussing the stochastic frontier models. In the case of cross-sectional models, it is usually assumed that the inefficiencies (ω_i) are i.i.d. with one-sided distribution such as truncated Gaussian (Aigner et al., 1977), exponential (Meeusen and van den Broeck, 1977) or gamma (Greene, 1990; Tsionas, 2000). Regarding panel data models, the inefficiency specification proposed by Battese and Coelli (1995) have been customarily used in empirical applications. Their panel data model proposal considers the Cobb Douglas production function (Cobb and Douglas, 1928) and can be written as

$$y_{it} = \mathbf{x}_{it}^{\top}\boldsymbol{\beta} + \varepsilon_{it} - \omega_{it}, \quad i = 1, \ldots, \mathrm{N}; \quad \text{and} \quad t = 1, \ldots, \mathrm{T}, \tag{6}$$

where y_{it} denotes the log-production for the i-th firm at time t, and the technical inefficiency ω_{it} is characterized to be truncated Gaussian distribution $\omega_{it} \sim \mathcal{N}_{[0,\infty)}(\mathbf{r}_{it}^{\top}\boldsymbol{\delta}, \sigma_{\omega}^2)$. Here, $\mathbf{r}_{it}$ is the vector of explanatory variables that explains the technical inefficiencies, $\boldsymbol{\delta}$ is the parameter vector associated to such variables and σ_{ω}^2 is a scale parameter. Moreover, an economic measure of efficiency - called *technical efficiency* (TE) - is defined as the ability to maximize the output from a given input quantity combination (Battese and Coelli, 1992) and it is computed as $\mathrm{TE}_{it} = \exp(-\omega_{it})$.

Concerning the use of Bayesian methods for inference in stochastic frontier models, van den Broeck et al. (1994) used Bayesian analysis based on importance sampling whilst

Tsionas (2002) resorted to Gibbs sampler to explore the posterior distribution in a model with random coefficients to separate technical inefficiency from technological differences across firms. In addition, Tsionas (2006) assumed a dynamic structure to model technical inefficiencies and another Gibbs sampling with data augmentation was employed.

More recently, Tsionas and Kumbhakar (2014) from a Bayesian approach, and Colombi et al. (2014); Filippini and Greene (2016); Lai and Kumbhakar (2018) from a classical approach considered a stochastic frontier panel data model in which the random firm-effects are separated from the persistent (time-invariant) and transient (time-varying) technical inefficiency.

3. Bayesian quantile stochastic frontier models

Let y_{it} be the output (the log-production, for example) and $\mathbf{x}_{it}$ be the input (the log-capital and the log-labour, for example) quantities for the i-th firm at time t, as defined in Section 2.2. Our proposal consists in exploring more widely the stochastic frontier through a quantile analysis of it. To simplify the exposition of our proposal, we assume that only two disturbance terms influence the productive process: the firm technical inefficiency ω_i and the error term ε_{it}. From here, our proposal can be extended to other structures.

For a fixed quantile $\tau \in (0,1)$, and conditioned to technical inefficiency $\omega_i^{(\tau)}$ (measure from the τ-quantile of stochastic frontier), let $f_\tau(\mathbf{x}_{it};\boldsymbol{\beta}_\tau)$ be the functional form which describes the relationship between the τ-quantile of output without inefficiency $y_{it}+\omega_i^{(\tau)}$ and input quantities. The parameter to be estimated, $\boldsymbol{\beta}_\tau$, and the technical inefficiency $\omega_i^{(\tau)}$ depends on the quantile τ (to simplify notation we will drop the τ index in the $\boldsymbol{\beta}_\tau$ and $\omega_i^{(\tau)}$ expressions hereinafter). Given theses assumptions, our proposed model can be written as

$$y_{it} = f_\tau(\mathbf{x}_{it};\boldsymbol{\beta}) + \varepsilon_{it} - \omega_i, \quad i = 1,\ldots,\mathrm{N}. \tag{7}$$

Various functional forms f can be specified. In particular, one can consider the linear functional form $f_\tau(\mathbf{x}_{it};\boldsymbol{\beta}) = \mathbf{x}_{it}^\top\boldsymbol{\beta}$, which is compatible, for example, with the Cobb Douglas and Translog production functions. Regarding the distribution specification of technical inefficiency, several distributions were considered in the literature, as discussed previously in Section 2.2. Here, we assume that the technical inefficiencies are i.i.d. with half-Gaussian distribution, $(\omega_i \mid \sigma_\omega^2) \sim \mathcal{N}_{[0,\infty)}(0, \sigma_\omega^2)$, where σ_ω^2 is a scale parameter. The random error distribution and the likelihood function are not specified.

As in Section 2.1 - to infer the model parameters from a Bayesian approach - we assume the asymmetric Laplace distribution with location parameter zero, asymmetric parameter τ (known and fixed) and unknown scale parameter σ_ε to the error term (Yu and Moyeed, 2001) which is denoted by $\mathcal{LA}_\tau(0; \sigma_\varepsilon)$. Through Gaussian-exponential mixture representation of the asymmetric Laplace distribution (Kotz et al., 2001), the model admits hierarchical representation and can be rewritten as

$$\begin{aligned} (y_{it} \mid \boldsymbol{\beta}, \sigma_\epsilon, \nu_{it}, \omega_i) &\sim \mathcal{N}(\mathbf{x}_{it}^\top \boldsymbol{\beta} + \kappa_1 \nu_{it} - \omega_i, \kappa_2 \sigma_\epsilon \nu_{it}) \\ (\nu_{it} \mid \sigma_\epsilon) &\sim \mathcal{E}(\sigma_\epsilon) \\ (\omega_i \mid \sigma_\omega^2) &\sim \mathcal{N}_{[0,\infty)}(0, \sigma_\omega^2), \end{aligned} \tag{8}$$

where $\kappa_1 = \dfrac{1-2\tau}{\tau(1-\tau)}$ and $\kappa_2 = \dfrac{2}{\tau(1-\tau)}$ are constants, $i = 1, \ldots, \mathrm{N}$ and $t = 1, \ldots, \mathrm{T}$.

To complete our Bayesian model specification, defining the parameter vector as $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_\varepsilon, \sigma_\omega^2)$, we take $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\sigma_\varepsilon)\pi(\sigma_\omega^2)$ such that $\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0)$, $\sigma_\varepsilon \sim \mathcal{IG}(n_{\varepsilon 0}, s_{\varepsilon 0})$ and $\sigma_\omega^2 \sim \mathcal{IG}(n_{\omega 0}, s_{\omega 0})$. This choice simplifies the computation of the full conditional distributions.

To simplify the notation, we have the following definitions: $\mathbf{y}_t = (y_{1t}, \ldots, y_{\mathrm{N}t})^\top$, $\mathbf{y} = (\mathbf{y}_1^\top, \ldots, \mathbf{y}_\mathrm{T}^\top)^\top$, $\mathbf{x}_t = (\mathbf{x}_{1t}, \ldots, \mathbf{x}_{\mathrm{N}t})^\top$, $\mathbf{X} = (\mathbf{x}_1^\top, \ldots, \mathbf{x}_\mathrm{N}^\top)^\top$, $\boldsymbol{\nu}_t = (\nu_{1t}, \ldots, \nu_{\mathrm{N}t})^\top$, $\boldsymbol{\nu} = (\boldsymbol{\nu}_1^\top, \ldots, \boldsymbol{\nu}_\mathrm{T}^\top)^\top$, $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)^\top$, and $\tilde{\boldsymbol{\omega}} = (\boldsymbol{\omega}^\top, \ldots, \boldsymbol{\omega}^\top)^\top$.

Note that $(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega}) \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{y}, \Sigma_\mathbf{y})$ with $\boldsymbol{\mu}_\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \kappa_1 \boldsymbol{\nu} - \tilde{\boldsymbol{\omega}}$ and $\Sigma_\mathbf{y} = \kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu})$, where $\mathrm{diag}(\boldsymbol{\nu})$ represents a diagonal matrix.

The posterior distributions $\pi(\boldsymbol{\theta} \mid \mathbf{y})$ and $\pi(\boldsymbol{\omega} \mid \mathbf{y})$ are obtained through a data augmentation scheme with latent variable $\boldsymbol{\nu}$, i.e. $\pi(\boldsymbol{\theta} \mid \mathbf{y}) = \int\int \pi(\boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega} \mid \mathbf{y}) d\boldsymbol{\nu} d\boldsymbol{\omega}$ and $\pi(\boldsymbol{\omega} \mid \mathbf{y}) = \int\int \pi(\boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega} \mid \mathbf{y}) d\boldsymbol{\nu} d\boldsymbol{\theta}$, where

$$\pi(\boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega} \mid \mathbf{y}) \propto \left[\prod\nolimits_{t,i} f(y_{it} \mid \boldsymbol{\beta}, \sigma_\varepsilon, \nu_{it}, \omega_i)\right] \left[\prod\nolimits_{t,i} f(\nu_{it} \mid \sigma_\varepsilon)\right] [\prod\nolimits_i f(\omega_i \mid \sigma_\omega^2)] \pi(\boldsymbol{\theta}).$$

The augmented posterior distribution $\pi(\boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega} \mid \mathbf{y})$ is not easy to manipulate algebraically.

However, full conditional distributions are available in known closed-form and easy to sample from. Such distributions are given by

$$\begin{aligned}
(\boldsymbol{\beta} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \sigma_\varepsilon) &\sim \mathcal{N}(\mathbf{m}, \mathbf{C}), \\
(\sigma_\varepsilon \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\beta}) &\sim \mathcal{IG}(n_\varepsilon, s_\varepsilon), \\
(\nu_{it} \mid y_{it}, \boldsymbol{\beta}, \sigma_\varepsilon, \omega_i) &\sim \mathcal{GIG}\left(0.5, \chi_{it}, \psi_{it}\right), \\
(\sigma_\omega^2 \mid \boldsymbol{\omega}) &\sim \mathcal{IG}(n_\omega, s_\omega), \\
(\omega_i \mid y_{it}, \nu_{it}, \boldsymbol{\beta}, \sigma_\varepsilon, \sigma_\omega^2) &\sim \mathcal{N}_{[0,\infty)}(m_{\omega_i}, C_{\omega_i}),
\end{aligned} \tag{9}$$

where

$$\begin{aligned}
\mathbf{C}^{-1} &= \mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1} \mathbf{X} + \mathbf{C}_0^{-1}, \\
\mathbf{m} &= \mathbf{C}\left[\mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1}(\mathbf{y} - \kappa_1 \boldsymbol{\nu} + \tilde{\boldsymbol{\omega}}) + \mathbf{C}_0^{-1}\mathbf{m}_0\right], \\
n_\varepsilon &= n_{\varepsilon 0} + 3\mathrm{NT}/2, \\
s_\varepsilon &= s_{\varepsilon 0} + \textstyle\sum_{t,i} \nu_{it} + (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\top [\mathrm{diag}(\boldsymbol{\nu})]^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})/(2\kappa_2), \\
\chi_{it} &= (y_{it} - \mathbf{x}_{it}^\top \boldsymbol{\beta} + \omega_i)^2/(\kappa_2 \sigma_\varepsilon), \\
\psi_{it} &= \left[2 + \kappa_1^2/\kappa_2\right]/\sigma_\varepsilon, \\
n_\omega &= n_{\omega 0} + \mathrm{N}/2, \\
s_\omega &= s_{\omega 0} + \boldsymbol{\omega}^\top \boldsymbol{\omega}/2, \\
C_{\omega_i}^{-1} &= (\kappa_2 \sigma_\epsilon)^{-1} \textstyle\sum_t \nu_{it}^{-1} + (\sigma_\omega^2)^{-1}, \\
m_{\omega_i} &= -C_{\omega_i} \textstyle\sum_t \left[(y_{it} - \mathbf{x}_{it}^\top \boldsymbol{\beta} - \kappa_1 \nu_{it})/(\kappa_2 \sigma_\varepsilon \nu_{it})\right].
\end{aligned}$$

The notation $\mathcal{GIG}(\cdot,\cdot,\cdot)$ denotes the *generalized inverse Gaussian* distribution (Jørgensen, 1982). The Gibbs sampling algorithm can be used to obtain a posterior sample of $\boldsymbol{\theta}$, $\boldsymbol{\nu}$ and $\boldsymbol{\omega}$.

Under regularity conditions, the (posterior) technical efficiency for the i-th firm, measured from the stochastic frontier τ-quantile and specified as $\mathrm{TE}_i = \mathbb{E}[\exp(-\omega_i)|\mathbf{y}]$, can be estimated from the posterior sample of ω_i by

$$\mathrm{TE}_i \approx \frac{1}{M} \textstyle\sum_k \exp(-\omega_i^{(k)}), \quad \omega_i^{(k)} \overset{iid}{\sim} \pi(\omega_i|\mathbf{y}), \quad k = 1, \ldots, M.$$

To construct credible intervals, for the quantile regression parameters ($\boldsymbol{\beta}$), with adequate credibility we can use the asymptotic posterior distribution of $\boldsymbol{\beta}$ that, as detailed in Yang et al. (2016), is approximately Gaussian. Let the asymptotic posterior distribution of $\boldsymbol{\beta}$ be $\mathcal{N}(\mu_{\boldsymbol{\beta}}, \Sigma_{\beta})$. The location parameter $\mu_{\boldsymbol{\beta}}$ can be estimated for the average of the posterior sample of $\boldsymbol{\beta}$ and the posterior covariance Σ_{β} is approximate as

$$\Sigma_{\boldsymbol{\beta}} \approx \frac{n\tau(1-\tau)}{\hat{\sigma}^2}\mathrm{Cov}(\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(M)})\mathbf{X}^{\top}\mathbf{X}\mathrm{Cov}(\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(M)}), \tag{10}$$

where $\hat{\sigma}^2$ denotes the point estimation of σ^2 (we use the average of the posterior sample), $\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(M)}$ denotes the posterior sample of $\boldsymbol{\beta}$ obtained from the use of Laplace asymmetric likelihood, and $\mathrm{Cov}(\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(M)})$ is the sample covariance matrix.

4. Monte Carlo study

Consider the model in Equation (7) with linear functional form, single covariate and intercept, i.e.

$$y_{it} = \beta_0 + \beta_1 x_{it} + \varepsilon_{it} - \omega_i, \quad i = 1, \ldots, \mathrm{N} \quad \text{and} \quad t = 1, \ldots, \mathrm{T}.$$

The functional form $\beta_0 + \beta_1 x$ represents the τ-quantile of the stochastic frontier. In this sense, the error term must satisfy $\Pr(\varepsilon_{it} \leqslant 0) = \tau$. Assuming that the error term distribution is Gaussian, then $\varepsilon_i \overset{iid}{\sim} \mathcal{N}(-\sigma_\varepsilon \Phi^{-1}(\tau), \sigma_\varepsilon^2)$, where $\Phi^{-1}(\cdot)$ denotes the inverse of the Gaussian cumulative distribution function.

Data sets are generated with the following settings: $\beta_0 = 0.3$, $\beta_1 = 0.5$, $\sigma_\varepsilon^2 = 0.01$, $x_{it} \overset{iid}{\sim} \mathcal{U}(0, 1)$, $\omega_i \overset{iid}{\sim} \mathcal{N}_{[0,\infty)}(0, \sigma_\omega^2 = 0.02)$, $\mathrm{N} = 30$ (number of firms), $\mathrm{T} \in \{20, 50, 100\}$ (number of periods) and $\tau \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$. An example from the generated data sets is exhibit in Figure 1 where the solid line represent the τ-quantile of the stochastic frontier ($\beta_0 + \beta_1 x_{it}$), the black points display the unobserved efficient production ($\beta_0 + \beta_1 x_{it} + \varepsilon_{it}$), and the gray points show the production y_{it}.

Each of the 15 cases (combination of the number of periods and quantiles) uses $M = 1,000$ Monte Carlo replications. A Gibbs sampler is implemented from the full conditional

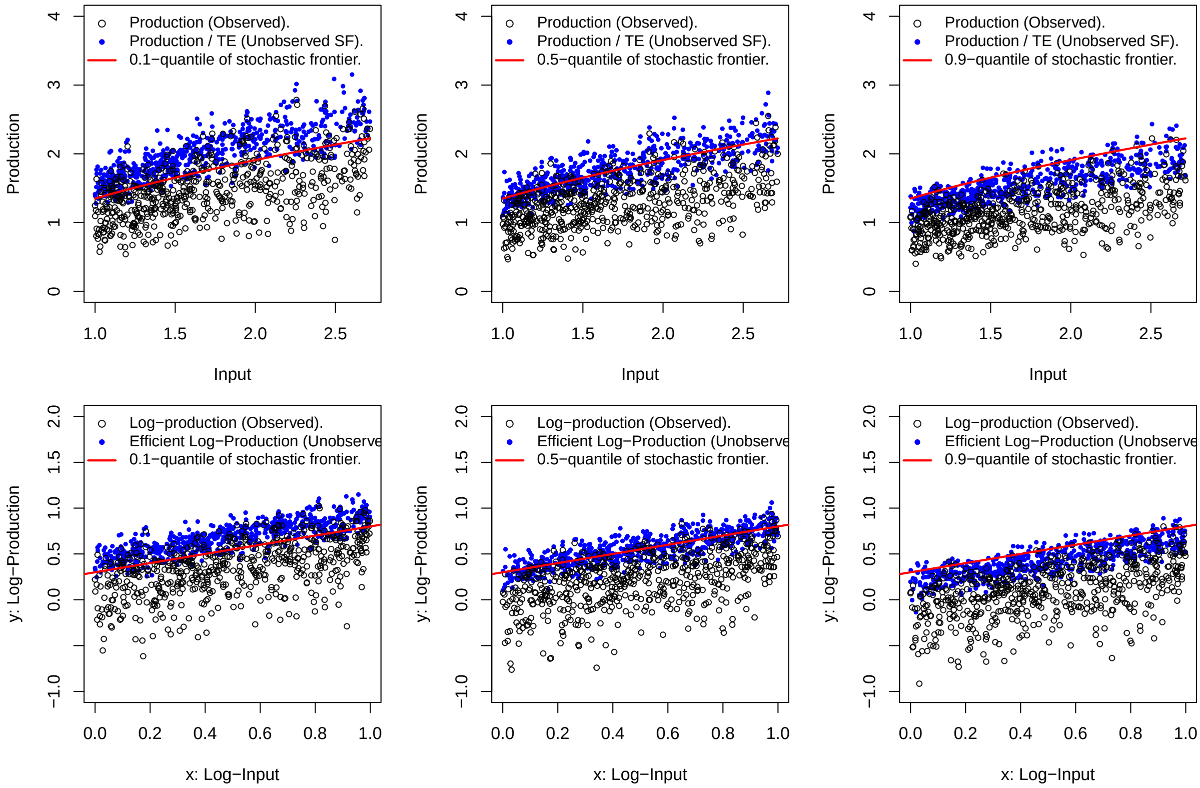

Figure 1: Data generate from $y_{it} = 0.3 + 0.5x_{it} + \varepsilon_{it} - \omega_i$, $x_{it} \overset{iid}{\sim} \mathcal{U}(0,1)$, $\varepsilon_{it} \overset{iid}{\sim} \mathcal{N}(0, 0.01)$ and $\omega_i \overset{iid}{\sim} \mathcal{N}_{[0,\infty)}(0, 0.02)$; for $i = 1, \ldots, 30$ and $t = 1, \ldots, 20$. The functional form $0.3 + 0.5x_{it}$, represented by the solid line, is the τ-quantile of stochastic frontier, for $\tau = 0.1$ (left), $\tau = 0.5$ (center) and $\tau = 0.9$ (right). The black points represent the unobserved efficient production and the gray points represent the (observed) production.

distributions in Equation (9). In preliminary tests, rapid convergence and low autocorrelations were observed in the generated MCMC chains. However, initial values for the technical efficiencies must be chosen with care. In our full study, we consider appropriate posterior samples obtained (in each replication) from two MCMC chains of length 5,100 with burn-in 99 and thinning 5.

We evaluate the consistency of the estimators through the (estimated) *mean squared error* (MSE), which is computed as

$$\text{MSE}(\theta) = \frac{1}{M}\sum_k \left[\hat{\theta}^{(k)} - \theta_0\right]^2,$$

where $\hat{\theta}^{(k)}$ is the posterior mean relative to the k-th Monte Carlo replication ($k = 1, \ldots, M$)

and θ_0 is the value of θ that was used to generates the observations.

Figure 2 summarises the MSE relative to the stochastic frontier parameters (the intercept on the left and the slope on the right) in each the five evaluate quantiles and each number of periods. The results suggest consistency of the parameter estimators.

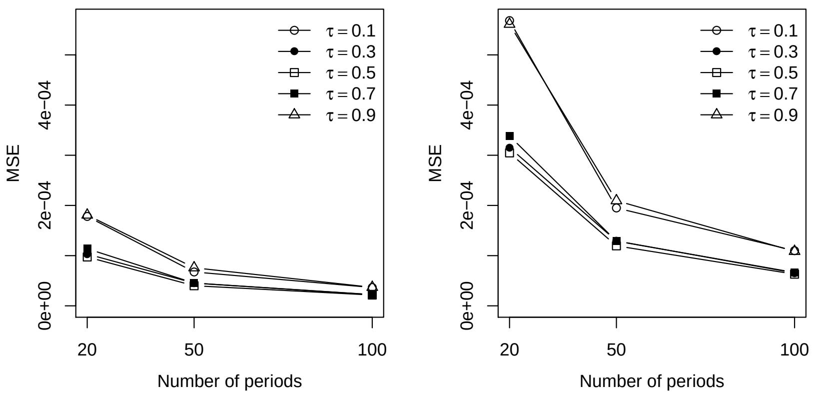

Figure 2: Mean squared error estimate of regression parameters: intercept (left) and slope (right).

Moreover, Figure 3 exhibits the MSE relative to the technical efficiency in each of the 30 firms in each the five evaluate quantiles and each number of periods. Here, the results suggest consistency of the technical efficiency estimators.

The credibility of the interval estimates are evaluated through *empirical coverage probability* (ECP), defined as

$$\mathrm{ECP}(\theta) = \frac{1}{M} \sum\nolimits_k \mathbb{1}_{\left[\theta^{(k)}_{\alpha/2}, \theta^{(k)}_{1-\alpha/2}\right]}(\theta_0),$$

where $\mathbb{1}_{\mathrm{A}}(\cdot)$ denotes the indicator function of a set A, $(1-\alpha)$ represent the credibility of the interval estimation to θ given by $\left[\theta^{(k)}_{\alpha/2}, \theta^{(k)}_{1-\alpha/2}\right]$ and $\theta^{(k)}_q$ denotes the estimated q-quantile of the posterior distribution of θ relative to the k-th Monte Carlo replication, with θ_0 defined as above. We consider $(1-\alpha) = 95\%$ credibility in all intervals computed.

When the limits of credible intervals (relative to the stochastic frontier parameters $\boldsymbol{\beta}$) are computed using the covariance of the Gaussian asymptotic distribution, showed in Equation (10), the ECP derived will be named *adjusted empirical coverage probability* and denoted by $\mathrm{ECP}_{\mathrm{adj}}$.

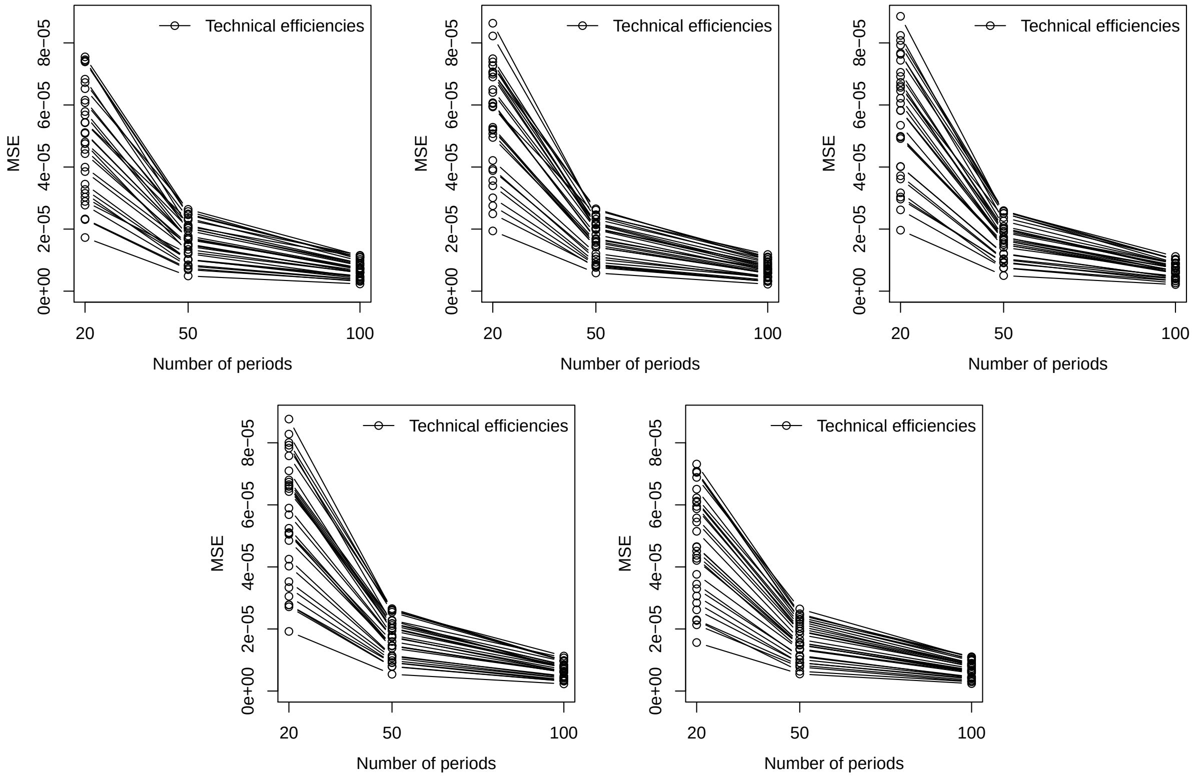

Figure 3: Mean squared error estimate of the technical efficiency in each of the 30 companies. Top: 0.1-quantile (left), 0.3-quantile (center) and 0.5-quantile (right). Bottom: 0.7-quantile (left) and 0.9-quantile (right).

Figure 4 exhibits the nominal 95% credibility (dotted line) and the ECP (in the top) or ECP_{adj} (in the bottom) of the stochastic frontier parameters $\boldsymbol{\beta}$ in each quantile for 20 (left), 50 (center) and 100 (right) periods. The results confirm that the employment of the asymptotic distribution is necessary for computing more adequate credible intervals for the stochastic frontier parameters, mainly when quantiles are away from the median.

The nominal 95% credibility and the ECP for the technical efficiencies of all 30 firms in the evaluated quantiles whilst the number of periods is 20 (left), 50 (center) and 100(right) are showed in Figure 5. This result shown an over-coverage in relation to the interval estimation of technical efficiencies.

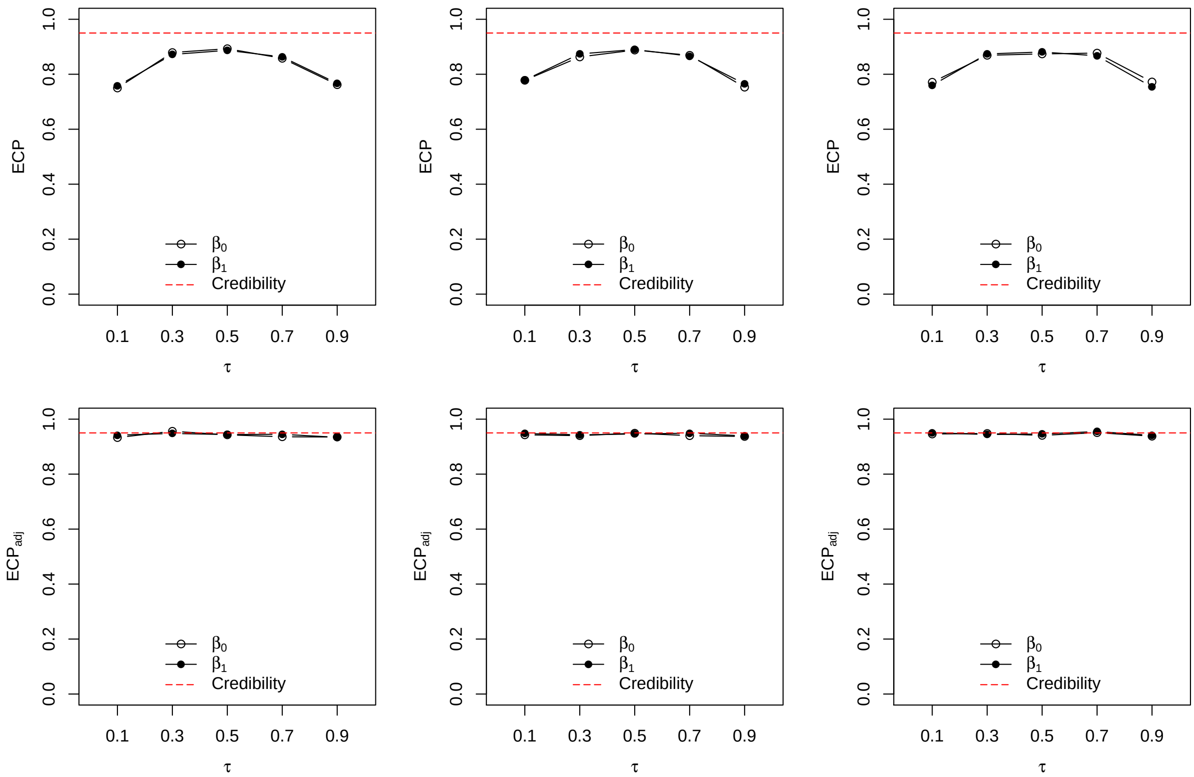

Figure 4: Nominal 95% credibility and ECP (top) or ECP_{adj} (bottom) of the stochastic frontier parameters for the five evaluated quantiles (τ) whilst the number of periods is 20 (left), 50 (center) and 100(right).

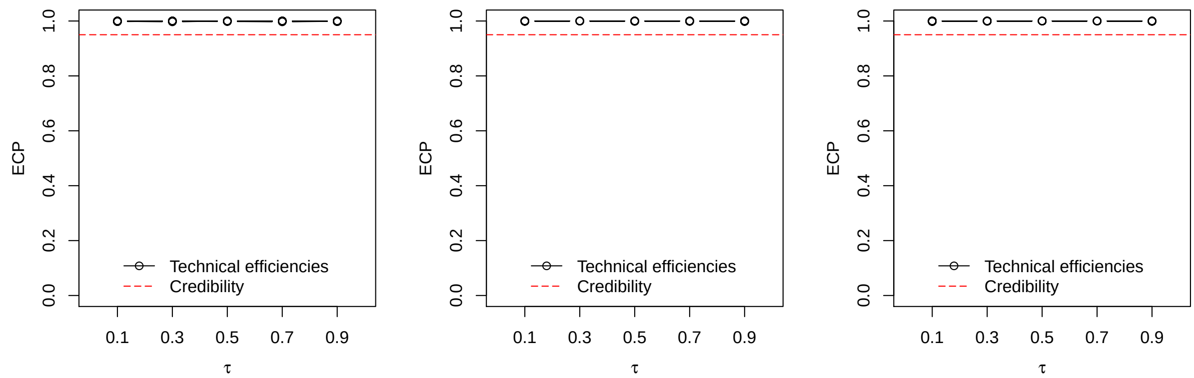

Figure 5: Nominal 95% credibility and ECP for all of the technical efficiencies for the five evaluated quantiles (τ) when the number of periods is 20 (left), 50 (center) and 100(right).

5. Applications

To illustrate the proposed model, we apply our methodology in two data sets. The first one, named `front41Data`, is a cross-sectional data of 60 firms. The second, named `riceProdPhil`, is a panel data with annual information collected from 43 smallholder rice producers in the Tarlac region of the Philippines between 1990 and 1997. Both data sets are available in the `frontier`[1] package (Coelli and Henningsen, 2013) of `R` (R Core Team, 2020).

5.1. *front41Data*

This data contain cross-sectional observations about *output* (production), *capital* and *labour* of 60 firms. The model can be written as

$$\log(output_i) = \beta_0 + \beta_1 \log(capital_i) + \beta_2 \log(labour_i) - \omega_i + \varepsilon_i, \quad i = 1, \ldots, 60,$$

where for the i-th firm, ω_i is the technical inefficiency and ε_i is the error term. The quantile regression parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^\top$ and technical inefficiencies depend of the quantile τ. The next assumptions are considered: $\omega_i \sim \mathcal{N}_{[0,\infty)}(0, \sigma^2_\omega)$, and, although the likelihood is misspecified, we use the asymmetric Laplace distribution to develop the Bayesian quantile regression in the stochastic production frontier, as detailed in Section 3.

Specifying a conjugate Gaussian-Gamma Inverse prior gives known full conditional distributions which, in turn, allow the implementation of the Gibbs sampling algorithm. Details about conditional distributions can be found in Appendix A (for T = 1). Each posterior sample (related to each quantile $\tau \in \{0.1, 0.2, \ldots, 0.9\}$) is composed of 1,000 elements drawn from 55,000 iterations, 5,000 as burn-in and 50 as thinning.

Results about the quantile regression parameters, $\boldsymbol{\beta}$, of the stochastic production frontier are exhibits in the Figure 6. The diagonal shown the posterior mean, and the 50% and 95% credible interval limits of $\boldsymbol{\beta}$ at each quantile, $\tau \in \{0.1, 0.2, \ldots, 0.9\}$, of the stochastic production frontier. The lower and upper triangle includes scatter plots related to posterior sample related to: the median ($\tau = 0.5$) regression parameters (lower triangle) and the

[1] The `frontier` package is used to obtain maximum likelihood estimates of stochastic frontier parameters, and estimates of mean and individual technical efficiencies. Two specifications are available: the error components specification with time-varying efficiencies (Battese and Coelli, 1992) and a model specification in which the firm effects are directly influenced by a number of variables (Battese and Coelli, 1995).

τ-quantile regression parameters - $\tau = 0.1$ (black) and $\tau = 0.9$ (gray) - (upper triangle), respectively.

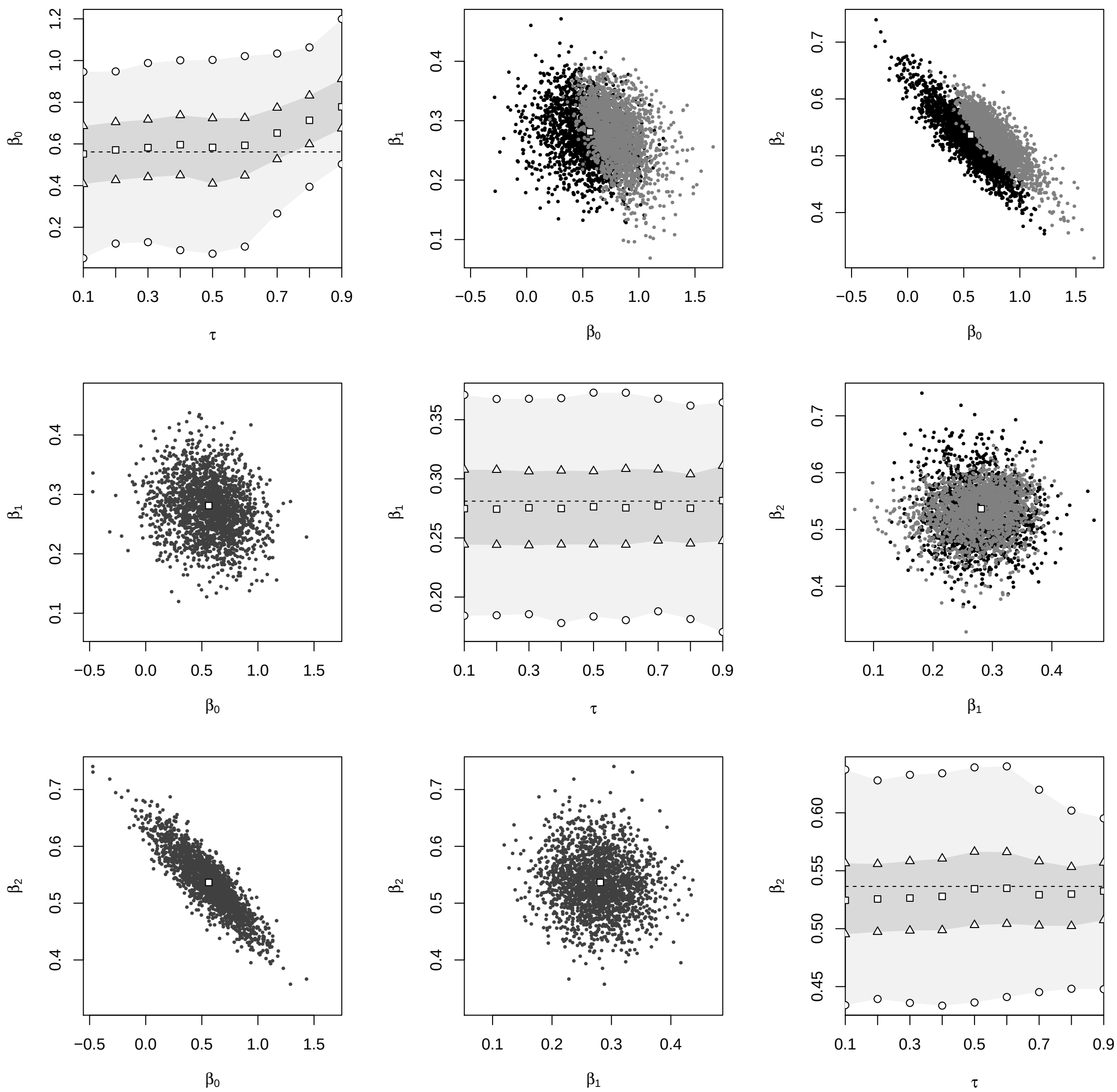

Figure 6: Posterior sample of the quantile regression parameters. Diagonal: Posterior mean (square), 50% credible interval limits (triangles) and 95% credible interval limits (circles) of the τ-quantile regression parameters, $\tau \in \{0.1, 0.2, \dots, 0.9\}$. Upper triangle: Scatter plots for the posterior sample of $\tau = 0.1$ (black) and $\tau = 0.9$ (gray) quantile regression parameters. Lower triangle: Scatter plots for the posterior sample of $\tau = 0.5$ quantile parameters. The dashed line (in the diagonal) and the white square (in the scatter plots) represent the maximum likelihood estimate provided for the `frontier` package.

The non-statistically significant results for the estimates corresponding to the log-capital (β_1) e log-labour (β_2) coefficients - with respect to the stochastic frontier quantiles - suggest presence of homocedasticity in the data. Regarding the intercept (β_0) the results indicate that the stochastic production frontier has low variability.

In this application our estimates are compatible with the maximum likelihood estimate due to homocedasticity and absence of outliers. Presence of heterocedasticity does not affect our proposed model, and outliers have less impact on the estimates compared with the fairly used conditional mean regression model. In summary, our proposal provides more complete information on the stochastic production frontier.

The technical efficiency is estimated for each firm $i = 1, \ldots, 60$ and for each analysed quantile of the stochastic frontier ($\tau \in \{0.1, 0.2, \ldots, 0.9\}$). The left and center plot in Figure 7 shown the credible intervals in each quantile of stochastic frontier of the firms 35 and 12 – firms with the lowest and highest technical efficiency respectively (we use the posterior medians to compare technical efficiencies).

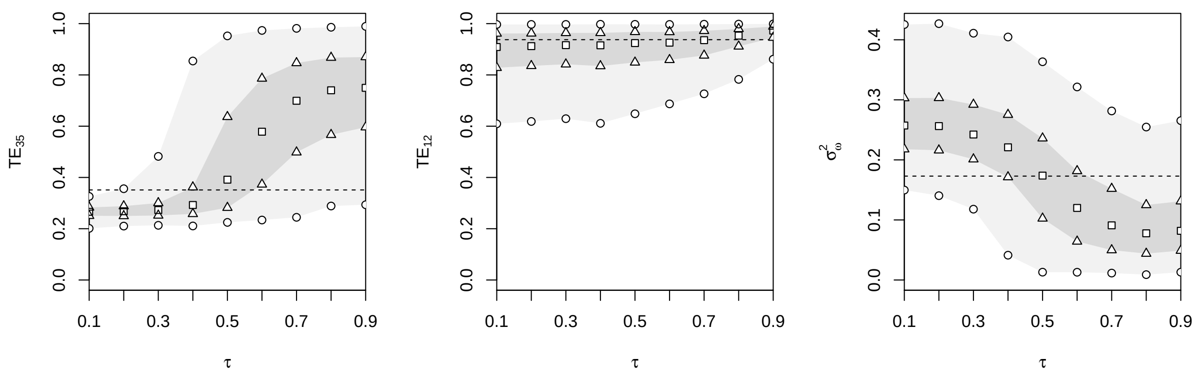

Figure 7: Posterior median (square), 0.025, 0.975 quantile (circle) and 0.25, 0.75 quantile (triangle) of the technical efficiencies, for the quantile regression of the stochastic production frontier was evaluate in the quantiles $\tau = 0.1, 0.2, \ldots, 0.9$, relative to the firm with smallest (left), and largest (center) technical efficiency mean. The dotted line represent the estimate obtained with the `frontier` package.

In this application we observed that, when measured from the lowest quantiles of the stochastic frontier, the technical efficiency of the firms reveals significant differences. This is corroborated in the Figure 8.

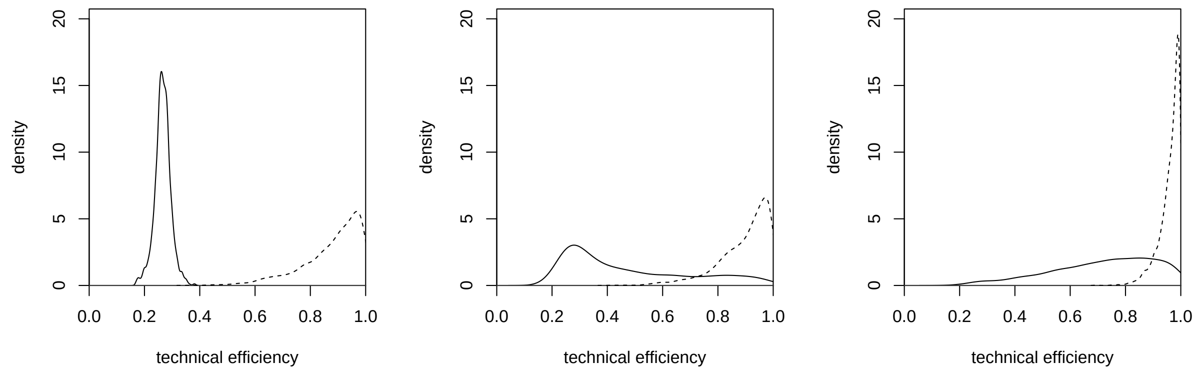

Figure 8: Posterior distribution of technical efficiencies related to firm 35 (continuous line) and 12 (dotted line), measured from the 0.1 (left), 0.5 (middle) and 0.9 (right) quantile of the stochastic frontier.

5.2. riceProdPhil

This data set, used as methodological illustration in several works (Coelli et al., 2005; Rho and Schmidt, 2015; Parmeter et al., 2019; Jradi et al., 2019, as example), corresponds to annual information over eight years of 43 rice producers. Details about data collection can be found in Pandey et al. (1999). For each rice producer and year, the information collected includes: *production* (tonnes of freshly threshed rice), *area* (hectares of planted rice), *labour* (man-days of family and hired labour) and *npk* (Fertiliser used, measured in Kg.).

Using the thanslog production function, the problem can be modeled as

$$\begin{aligned}\log(production_{it}) &= \beta_0 + \beta_1 \log(area_{it}) + \beta_2 \log(labour_{it}) + \beta_3 \log(npk_{it}) \\ &\quad + \beta_4 \log(area_{it})^2/2 + \beta_5 \log(labour_{it})^2/2 + \beta_6 \log(npk_{it})^2/2 \\ &\quad + \beta_7 \log(area_{it}) \log(labour_{it}) + \beta_8 \log(area_{it}) \log(npk_{it}) \\ &\quad + \beta_9 \log(labour_{it}) \log(npk_{it}) + \varepsilon_{it} - \omega_i\end{aligned}$$

where, for the i-th firm at time t, ω_i denotes the technical inefficiency and ε_{it} is the error term. It is important to remember that the quantile regression parameters $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_9)^\top$ and the technical inefficiencies depend of the analyzed quantile (τ) of the stochastic frontier. We assume inefficiencies with half-normal distribution, i.e. $\omega_i \sim \mathcal{N}_{[0,\infty)}(0, \sigma^2_\omega)$, and again although the likelihood is misspecified, we assume errors with asymmetric Laplace distribution $\varepsilon_{it} \sim \mathcal{LA}_\tau(0, \sigma_\varepsilon)$ to infer the parameters and technical (in)efficiencies using the proposal presented in Section 3.

To complete the specification we consider the prior $\pi(\boldsymbol{\beta}, \sigma_\varepsilon, \sigma^2_\omega) = \pi(\boldsymbol{\beta})\pi(\sigma_\varepsilon)\pi(\sigma^2_\omega)$, where $\beta \sim \mathcal{N}(0, 100)$, $\sigma_\varepsilon \sim \mathcal{IG}(0.01, 0.01)$ and $\sigma^2_\omega \sim \mathcal{IG}(0.01, 0.01)$.

The posterior sample for each regression (one for each quantile evaluate) was obtained from two MCMC chains that converged quickly and had low autocorrelations, in each chain with 1,000 elements drawn from 5,100 iterations, 99 as burn-in and 5 as thinning.

Results are exhibited in Figure 9 which shows for each quantile regression parameter (β_p, $p = 0, \ldots, 9$), the median (square), 50% credible interval limits (triangles) and 95% credible interval limits (circles) in each evaluated quantile ($\tau \in \{0.1, 0.2, 0.3, 0.5, 0.7, 0.9\}$) of the stochastic frontier. The limits of $(1 - \alpha)\%$ credible interval are given by the quantiles $\alpha/2$ and $1 - \alpha/2$ of the posterior sample.

Regarding the intercept (β_0) the results show that the posterior median increases when the analyzed quantile of the stochastic frontier is higher. This is consistent with the assumption of the linearity of the relationship between the inputs and stochastic frontier quantiles. Regarding the other regression parameters (β_p, $p = 1, \ldots, 9$), we observed slight differences among credible intervals for differing quantiles of the stochastic frontier evaluated. This may imply heteroskedasticity in the data, although in some cases these differences may be non-statistically significant. Thus, it suggests a possibly more flexible modeling such as median regression or other specifications.

In particular, we can focus on the median regression - a most robust regression compared with the conditional mean. The Figure 10 shows the posterior densities (dashed lines), the asymptotic distributions (solid line) and the prior densities (dotted line) of the median regression parameters in the `riceProdPhil` dataset.

The posterior median and the 95% credible intervals to technical efficiencies relative to median regression are showed in Figure 11. This results allow us to compare the performance of the firms productivity and to find significantly statistical differences.

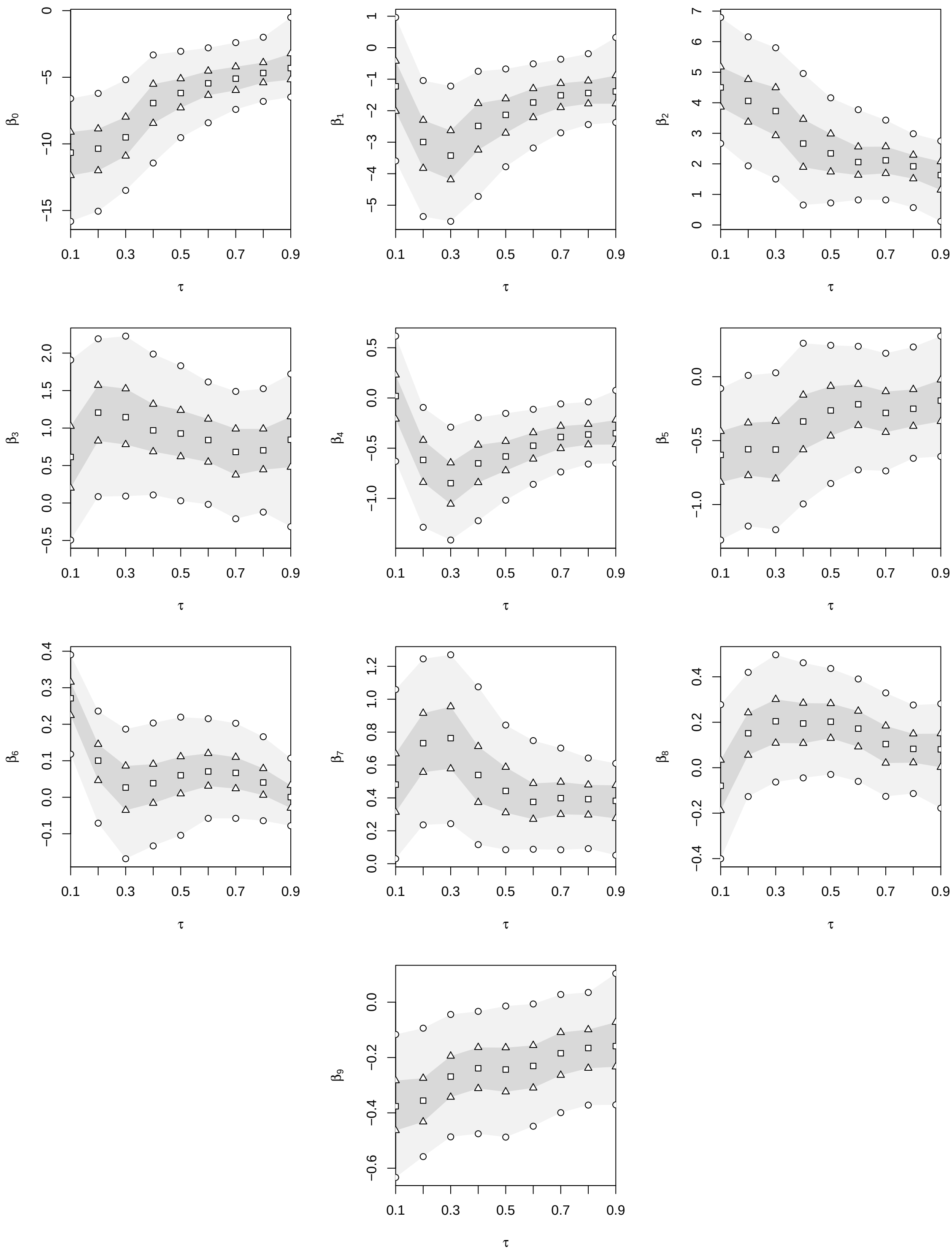

Figure 9: Posterior sample related to `riceProdPhil` data.

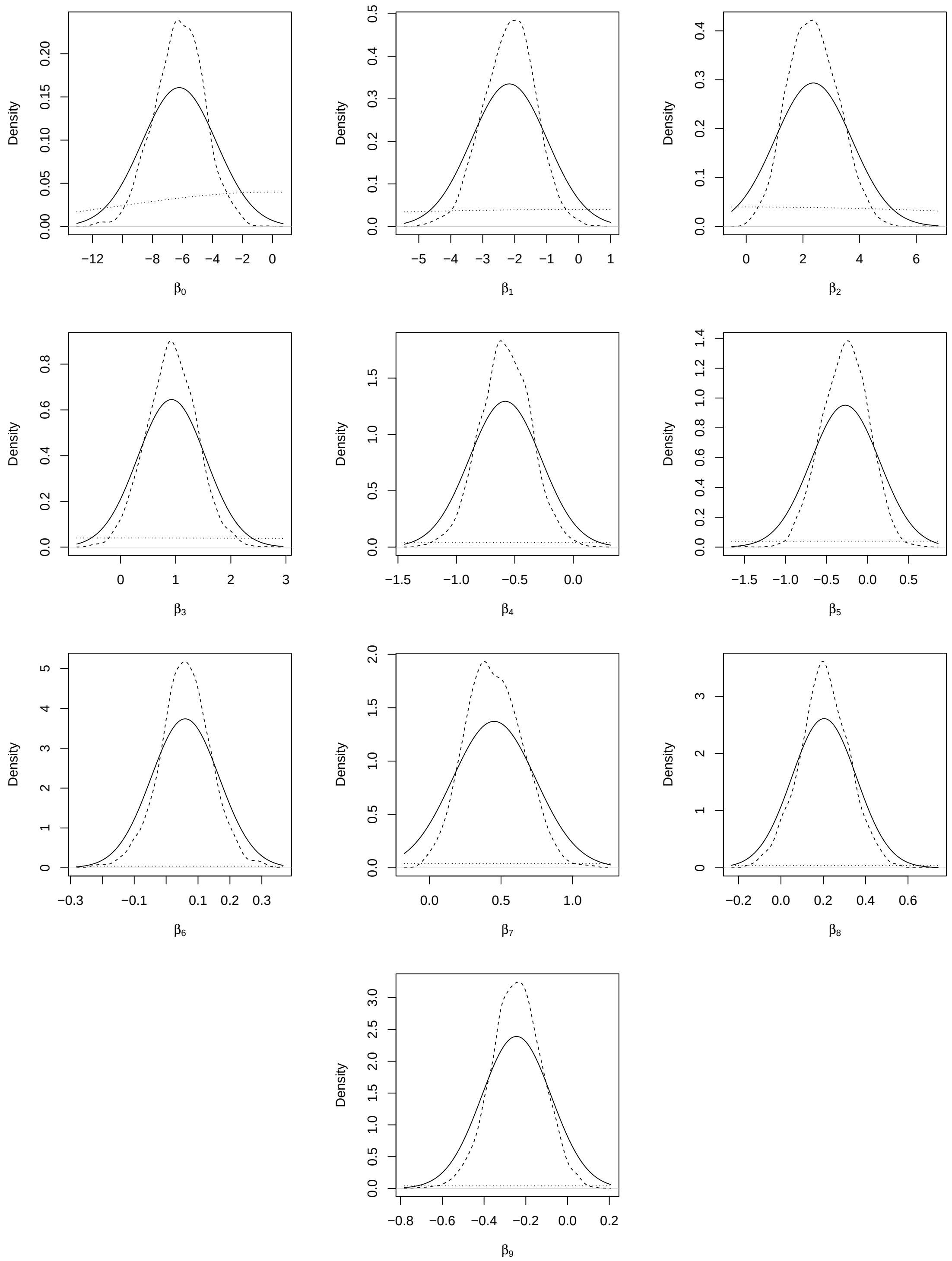

Figure 10: Asymptotic (solid line), posterior (dashed line) and prior (dotted line) densities related to the median regression parameters in `riceProdPhil` data.

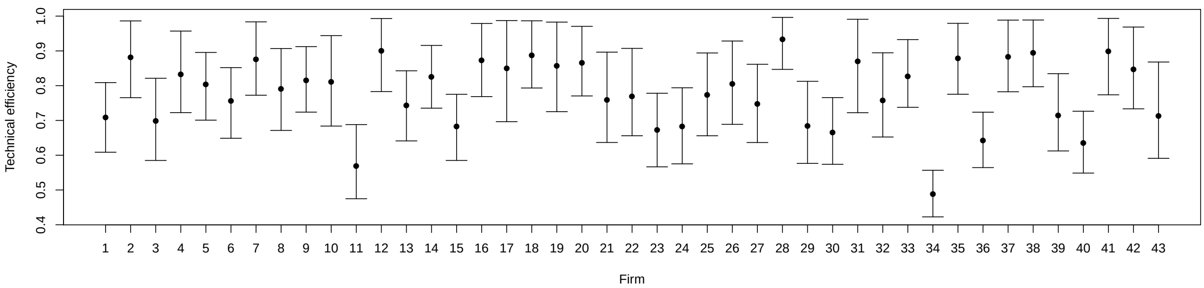

Figure 11: Posterior median and the 95% credible intervals to technical efficiencies for the median regression stochastic frontier model.

6. Concluding remarks

A quantile exploration of the relationship between inputs and outputs in stochastic frontier models can be beneficial in many senses. First, it is possible - since the evaluation of several stochastic frontier quantiles - to form an opinion about the adequacy of the functional form of the production function and presence or not of homocedasticity in the data. The technical efficiency can be measured from various stochastic frontier quantiles. In particular, when the stochastic frontier quantile is the median, we have a robust alternative to the conditional mean approach which deals with heteroscedasticity and does not need a specification of the likelihood.

Our proposal provides a method to infer the quantile regression parameters and technical efficiencies from a Bayesian perspective through a fast converging Gibbs sampler. Although the estimated covariance matrix for the parameters is underestimated, it is possible to adjust such covariance matrix as showed in Section 3.

Our proposed model can be extended in several directions. It can include time-varying stochastic frontier parameter (Gonçalves et al., 2020), or explore the pertinence of stochastic frontier of four factors. One may also study other distributions for technical inefficiency.

References

Aigner, D., Lovell, C.A.K., Schmidt, P., 1977. Formulation and estimation of stochastic frontier production function models. Journal of Econometrics 6, 21–37. doi:10.1016/0304-4076(77)90052-5.

Battese, G.E., Coelli, T.J., 1992. Frontier production functions, technical efficiency and panel data: With application to paddy farmers in India. Journal of Productivity Analysis 3, 153–169. doi:10.1007/BF00158774.

Battese, G.E., Coelli, T.J., 1995. A model for technical inefficiency effects in a stochastic frontier production function for panel data. Empirical Economics 20, 325–332. doi:10.1007/BF01205442.

Behr, A., 2010. Quantile regression for robust bank efficiency score estimation. European Journal of Operational Research 200, 568–581. doi:10.1016/j.ejor.2008.12.033.

Bernini, C., Freo, M., Gardini, A., 2004. Quantile estimation of frontier production function. Empirical Economics 29, 373–381. doi:10.1007/s00181-003-0173-5.

van den Broeck, J., Koop, G., Osiewalski, J., Steel, M.F.J., 1994. Stochastic frontier models: A Bayesian perspective. Journal of Econometrics 61, 273–303. doi:10.1016/0304-4076(94)90087-6.

Cobb, C.W., Douglas, P.H., 1928. A theory of production. The American Economic Review 18, 139–165.

Coelli, T., Henningsen, A., 2013. frontier: Stochastic Frontier Analysis. URL: `http://cran.r-project.org/package=frontier`.

Coelli, T.J., Rao, D.S.P., O'Donnell, C.J., Battese, G.E., 2005. An Introduction to Efficiency and Productivity Analysis. 2nd ed., Springer.

Colombi, R., Kumbhakar, S.C., Martini, G., Vittadini, G., 2014. Closed-skew normality in stochastic frontiers with individual effects and long/short-run efficiency. Journal of Productivity Analysis 42, 123–136. doi:10.1007/s11123-014-0386-y.

Filippini, M., Greene, W.H., 2016. Persistent and transient productive inefficiency: A maximum simulated likelihood approach. Journal of Productivity Analysis 45, 187–196. doi:10.1007/s11123-015-0446-y.

Gonçalves, K.C.M., Migon, H.S., Bastos, L.S., 2020. Dynamic quantile linear models: A Bayesian approach. Bayesian Analysis 15, 335–362. doi:10.1214/19-BA1156.

Greene, W.H., 1990. A Gamma-distributed stochastic frontier model. Journal of Econometrics 46, 141–163. doi:10.1016/0304-4076(90)90052-U.

Griffin, J.E., Steel, M.F.J., 2007. Bayesian stochastic frontier analysis using WinBUGS. Journal of Productivity Analysis 27, 163–176. doi:10.1007/s11123-007-0033-y.

Jørgensen, B., 1982. Statistical Properties of the Generalized Inverse Gaussian Distribution. volume 9 of *Lecture Notes in Statistics*. Springer. doi:10.1007/978-1-4612-5698-4.

Jradi, S., Parmeter, C.F., Ruggiero, J., 2019. Quantile estimation of the stochastic frontier model. Economics Letters 182, 15–18. doi:10.1016/j.econlet.2019.05.038.

Koenker, R., 2005. Quantile Regression. volume 38 of *Econometric Society Monographs*. Cambridge University Press. doi:10.1017/CBO9780511754098.

Kotz, S., Kozubowski, T.J., Podgórski, K., 2001. The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance. Birkhäuser Basel. doi:10.1007/978-1-4612-0173-1.

Kozumi, H., Kobayashi, G., 2011. Gibbs sampling methods for Bayesian quantile regression. Journal of Statistical Computation and Simulation 81, 1565–1578. doi:10.1080/00949655.2010.496117.

Lai, H., Kumbhakar, S.C., 2018. Panel data stochastic frontier model with determinants of persistent and transient inefficiency. European Journal of Operational Research 271, 746–755. doi:10.1016/j.ejor.2018.04.043.

Li, M., 2015. Moving beyond the linear regression model. Journal of Management 41, 71–98. doi:10.1177/0149206314551963.

Liu, C., Laporte, A., Ferguson, B.S., 2008. The quantile regression approach to efficiency measurement: Insights from Monte Carlo simulations. Health Economics 17, 1073–1087. doi:10.1002/hec.1398.

Meeusen, W., van den Broeck, J., 1977. Technical efficiency and dimension of the firm: Some results on the use of frontier production functions. Empirical Economics 2, 109–122. doi:10.1007/BF01767476.

Pandey, S., Masicat, P., Velasco, L., Villano, R.A., 1999. Risk analysis of a rainfed rice production system in Tarlac, Central Luzon, Philippines. Experimental Agriculture 35, 225–237. doi:10.1017/S0014479799002021.

Parmeter, C.F., Wan, A.T.K., Zhang, X., 2019. Model averaging estimators for the stochastic frontier model. Journal of Productivity Analysis 51, 91–103. doi:10.1007/s11123-019-00547-8.

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: `http://www.r-project.org/`.

Rho, S., Schmidt, P., 2015. Are all firms inefficient? Journal of Productivity Analysis 43, 327–349. doi:10.1007/s11123-013-0374-7.

Sriram, K., Ramamoorthi, R.V., Ghosh, P., 2013. Posterior consistency of Bayesian quantile regression based on the misspecified asymmetric Laplace density. Bayesian Analysis 8, 479–504. doi:10.1214/13-BA817.

Tsionas, E.G., 2000. Full likelihood inference in Normal-Gamma stochastic frontier models. Journal of Productivity Analysis 13, 183–205. doi:10.1023/A:1007845424552.

Tsionas, E.G., 2002. Stochastic frontier models with random coefficients. Journal of Applied Econometrics 17, 127–147. doi:10.1002/jae.637.

Tsionas, E.G., 2006. Inference in dynamic stochastic frontier models. Journal of Applied Econometrics 21, 669–676. doi:10.1002/jae.862.

Tsionas, E.G., Kumbhakar, S.C., 2014. Firm heterogeneity, persistent and transient technical inefficiency: A generalized true random-effects models. Journal of Applied Econometrics 29, 110–132. doi:10.1002/jae.2300.

Yang, Y., Wang, H.J., He, X., 2016. Posterior inference in Bayesian quantile regression with asymmetric

Laplace likelihood. International Statistical Review 84, 327–344. doi:10.1111/insr.12114.
Yu, K., Moyeed, R., 2001. Bayesian quantile regression. Statistics & Probability Letters 54, 437–447. doi:10.1016/S0167-7152(01)00124-9.

Appendix A. Full conditional distributions

The rewritten model in Equation (8) is given by

$$\begin{aligned}(y_{it} \mid \boldsymbol{\beta}, \sigma_\epsilon, \nu_{it}, \omega_i) &\sim \mathcal{N}(\mathbf{x}_{it}^\top \boldsymbol{\beta} + \kappa_1 \nu_{it} - \omega_i, \kappa_2 \sigma_\epsilon \nu_{it}) \\ (\nu_{it} \mid \sigma_\epsilon) &\sim \mathcal{E}(\sigma_\epsilon) \\ (\omega_i \mid \sigma_\omega^2) &\sim \mathcal{N}_{[0,\infty)}(0, \sigma_\omega^2), \quad i = 1, \ldots, \mathrm{N} \quad \text{and} \quad t = 1, \ldots, \mathrm{T}.\end{aligned}$$

where $\kappa_1 = [1 - 2\tau]/[\tau(1-\tau)]$ and $\kappa_2 = 2/[\tau(1-\tau)]$ are constants.

We specified the prior distribution to $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma_\varepsilon, \sigma_\omega^2)$ as $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\sigma_\varepsilon)\pi(\sigma_\omega^2)$, such that: $\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{m}_{\boldsymbol{\beta}0}, \mathbf{C}_{\boldsymbol{\beta}0})$; $\sigma_\varepsilon \sim \mathcal{IG}(n_{\varepsilon 0}, s_{\varepsilon 0})$ and $\sigma_\omega^2 \sim \mathcal{IG}(n_{\omega 0}, s_{\omega 0})$.

Using the Bayes theorem, the augmented posterior distribution to the parameters $\boldsymbol{\theta}$, auxiliary variable and technical inefficiencies is proportional to

$$\left[\prod_{t,i} f(y_{it} \mid \boldsymbol{\beta}, \sigma_\varepsilon, \nu_{it}, \omega_i)\right] \left[\prod_{t,i} f(\nu_{it} \mid \sigma_\varepsilon)\right] \left[\prod_i f(\omega_i \mid \sigma_\omega^2)\right] \pi(\boldsymbol{\theta}),$$

since that the prior are a proper distribution.

To simplify the notation we define the next elements for $t = 1, \ldots, \mathrm{T}$: $\mathbf{y}_t = (y_{1t}, \ldots, y_{\mathrm{N}t})^\top$, $\mathbf{y} = (\mathbf{y}_1^\top, \ldots, \mathbf{y}_\mathrm{T}^\top)^\top$, $\mathbf{x}_t = (\mathbf{x}_{1t}, \ldots, \mathbf{x}_{\mathrm{N}t})^\top$, $\mathbf{X} = (\mathbf{x}_1^\top, \ldots, \mathbf{x}_\mathrm{N}^\top)^\top$, $\boldsymbol{\nu}_t = (\nu_{1t}, \ldots, \nu_{\mathrm{N}t})^\top$, $\boldsymbol{\nu} = (\boldsymbol{\nu}_1^\top, \ldots, \boldsymbol{\nu}_\mathrm{T}^\top)^\top$, $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_N)^\top$ and the $(\mathrm{N} \times \mathrm{T})$-dimensional vector $\tilde{\boldsymbol{\omega}} = (\boldsymbol{\omega}^\top, \ldots, \boldsymbol{\omega}^\top)^\top$.

The calculation of full conditional distributions is displayed below (all constant terms were ignored).

Note that $(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega}) \sim \mathcal{N}(\boldsymbol{\mu}_\mathbf{y}, \Sigma_\mathbf{y})$ with $\boldsymbol{\mu}_\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \kappa_1 \boldsymbol{\nu} - \tilde{\boldsymbol{\omega}}$ and $\Sigma_\mathbf{y} = \kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu})$, being that $\mathrm{diag}(\boldsymbol{\nu})$ denotes a matrix with diagonal given by $\boldsymbol{\nu}$ and zero off-diagonal entries.

Thus

$$\begin{aligned}
\pi(\boldsymbol{\beta} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \sigma_\varepsilon) &\propto \left[\prod_{t,i} f(y_{it} \mid \boldsymbol{\beta}, \sigma_\varepsilon, \nu_{it}, \omega_i)\right] \pi(\boldsymbol{\beta}) \\
&\propto \exp\left[-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\top \Sigma_{\mathbf{y}}^{-1}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})\right] \exp\left[-\frac{1}{2}(\boldsymbol{\beta} - \mathbf{m}_0)^\top \mathbf{C}_0^{-1}(\boldsymbol{\beta} - \mathbf{m}_0)\right] \\
&\propto \exp\left[-\frac{1}{2}(\boldsymbol{\beta}^\top \mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1} \mathbf{X}\boldsymbol{\beta} - \boldsymbol{\beta}^\top \mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1}(\mathbf{y} - \kappa_1 \boldsymbol{\nu} + \tilde{\boldsymbol{\omega}})\right. \\
&\quad -(\mathbf{y} - \kappa_1 \boldsymbol{\nu} + \tilde{\boldsymbol{\omega}})^\top \Sigma_{\mathbf{y}}^{-1} \mathbf{X}\boldsymbol{\beta})] \\
&\quad \exp\left[-\frac{1}{2}(\boldsymbol{\beta}^\top \mathbf{C}_0^{-1}\boldsymbol{\beta} - \boldsymbol{\beta}^\top \mathbf{C}_0^{-1}\mathbf{m}_0 - \mathbf{m}_0^\top \mathbf{C}_0^{-1}\boldsymbol{\beta})\right]
\end{aligned}$$

Define $\mathbf{C}^{-1} = \mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1}\mathbf{X} + \mathbf{C}_0^{-1}$, then

$$\begin{aligned}
\pi(\boldsymbol{\beta} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \sigma_\varepsilon) &\propto \exp\left[-\frac{1}{2}(\boldsymbol{\beta}^\top \mathbf{C}^{-1}\boldsymbol{\beta} - \boldsymbol{\beta}^\top \mathbf{C}^{-1}\mathbf{C}[\mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1}(\mathbf{y} - \kappa_1 \boldsymbol{\nu} + \tilde{\boldsymbol{\omega}}) + \mathbf{C}_0^{-1}\mathbf{m}_0]\right. \\
&\quad \left. -[\mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1}(\mathbf{y} - \kappa_1 \boldsymbol{\nu} + \tilde{\boldsymbol{\omega}}) + \mathbf{C}_0^{-1}\mathbf{m}_0]^\top \mathbf{C}^\top \mathbf{C}^{-1}\boldsymbol{\beta})\right],
\end{aligned}$$

define also $\mathbf{m} = \mathbf{C}\left[\mathbf{X}^\top \Sigma_{\mathbf{y}}^{-1}(\mathbf{y} - \kappa_1 \boldsymbol{\nu} + \tilde{\boldsymbol{\omega}}) + \mathbf{C}_0^{-1}\mathbf{m}_0\right]$, thus

$$\pi(\boldsymbol{\beta} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \sigma_\varepsilon) \propto \exp\left[-\frac{1}{2}(\boldsymbol{\beta} - \mathbf{m})^\top \mathbf{C}^{-1}(\boldsymbol{\beta} - \mathbf{m})\right],$$

which is the kernel of the Gaussian distribution, i.e. $(\boldsymbol{\beta} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \sigma_\varepsilon) \sim \mathcal{N}(\mathbf{m}, \mathbf{C})$.

$$\begin{aligned}
\pi(\sigma_\varepsilon \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\beta}) &\propto f(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega})\left[\prod_{t,i} f(\nu_{it} \mid \sigma_\varepsilon)\right]\pi(\sigma_\varepsilon) \\
&\propto |\Sigma_{\mathbf{y}}|^{-1/2} \exp\left[-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\top \Sigma_{\mathbf{y}}^{-1}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})\right]\left[\prod_{t,i} \sigma_\varepsilon \exp\left(-\frac{\nu_{it}}{\sigma_\varepsilon}\right)\right] \\
&\quad \sigma_\varepsilon^{-(n_{\varepsilon 0}+1)} \exp\left[-\frac{s_{\varepsilon 0}}{\sigma_\varepsilon}\right] \\
&\propto \sigma_\varepsilon^{-(n_{\varepsilon 0}+3\mathrm{NT}/2+1)} \\
&\times \exp\left[-\frac{1}{\sigma_\varepsilon}\left(s_{\varepsilon 0} + \sum_{i,t} \nu_{it} + \frac{(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\top [\mathrm{diag}(\boldsymbol{\nu})]^{-1}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})}{2\kappa_2}\right)\right]
\end{aligned}$$

which is a kernel of the inverse gamma distribution, i.e. $(\sigma_\varepsilon \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\omega}, \boldsymbol{\beta}) \sim \mathcal{IG}(n_\varepsilon, s_\varepsilon)$ where

$$n_\varepsilon = n_{\varepsilon 0} + 3\mathrm{NT}/2 \quad \text{and} \quad s_\varepsilon = s_{\varepsilon 0} + \sum_{it} \nu_{it} + (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\top [\mathrm{diag}(\boldsymbol{\nu})]^{-1}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})/(2\kappa_2).$$

$$
\begin{aligned}
\pi(\nu_{it} \mid y_{it}, \omega_i, \boldsymbol{\beta}, \sigma_\varepsilon) &\propto f(y_{it} \mid \boldsymbol{\theta}, \nu_{it}, \omega_i) f(\nu_{it} \mid \sigma_\varepsilon) \\
&\propto (\kappa_2 \sigma_\varepsilon \nu_{it})^{-1/2} \exp\left[-\frac{1}{2}\frac{(y_{it} - \mathbf{x}_{it}^\top \boldsymbol{\beta} - \kappa_1 \nu_{it} + \omega_i)^2}{\kappa_2 \sigma_\varepsilon \nu_{it}}\right] \sigma_\varepsilon \exp\left(-\frac{\nu_{it}}{\sigma_\varepsilon}\right) \\
&\propto {\nu_{it}}^{-1/2} \exp\left[-\frac{1}{2}\left(\left[\frac{\kappa_1^2}{\kappa_2 \sigma_\varepsilon} + \frac{2}{\sigma_\varepsilon}\right] \nu_{it} + \frac{(y_{it} - \mathbf{x}_{it}^\top \boldsymbol{\beta} + \omega_i)^2}{\kappa_2 \sigma_\varepsilon} \frac{1}{\nu_{it}}\right)\right]
\end{aligned}
$$

which is a kernel of the generalized inverse Gaussian distribution, denoted by $(\nu_{it} \mid y_{it}, \omega_i, \boldsymbol{\beta}, \sigma_\varepsilon) \sim \mathcal{GIG}(1/2, \chi_{it}, \psi_{it})$ where

$$
\chi_{it} = \frac{(y_{it} - \mathbf{x}_{it}^\top \boldsymbol{\beta} + \omega_i)^2}{\kappa_2 \sigma_\varepsilon} \quad \text{and} \quad \psi_{it} = \left[\frac{\kappa_1^2}{\kappa_2} + 2\right] \frac{1}{\sigma_\varepsilon}.
$$

$$
\begin{aligned}
\pi(\sigma_\omega^2 \mid \boldsymbol{\omega}) &\propto \left[\textstyle\prod_i f(\omega_i \mid \sigma_\omega^2)\right] \pi(\sigma_\omega^2) \\
&\propto (\sigma_\omega^2)^{-\mathrm{N}/2} \exp\left[-\frac{1}{2\sigma_\omega^2} \boldsymbol{\omega}^\top \boldsymbol{\omega}\right] (\sigma_\omega^2)^{-(n_{\omega 0}+1)} \exp\left[-\frac{s_{\omega 0}}{\sigma_\omega^2}\right] \\
&\propto (\sigma_\omega^2)^{-(n_{\omega 0}+\mathrm{N}/2+1)} \exp\left[-\frac{1}{\sigma_\omega^2}\left(s_{\omega 0} + \frac{\boldsymbol{\omega}^\top \boldsymbol{\omega}}{2}\right)\right]
\end{aligned}
$$

which is a kernel of the inverse gamma distribution, i.e. $(\sigma_\omega^2 \mid \boldsymbol{\omega}) \sim \mathcal{IG}(n_\omega, s_\omega)$ where

$$
n_\omega = n_{\omega 0} + \mathrm{N}/2 \quad \text{and} \quad s_\omega = s_{\omega 0} + \boldsymbol{\omega}^\top \boldsymbol{\omega}/2.
$$

Let $\mathbb{I}_A(\cdot)$ be the indicator function of a subset A, then

$$
\begin{aligned}
\pi(\boldsymbol{\omega} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\theta}) &\propto f(\mathbf{y} \mid \boldsymbol{\theta}, \boldsymbol{\nu}, \boldsymbol{\omega}) f(\boldsymbol{\omega} \mid \sigma_\omega^2) \\
&\propto \exp\left[-\frac{1}{2}(\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})^\top \Sigma_{\mathbf{y}}^{-1} (\mathbf{y} - \boldsymbol{\mu}_{\mathbf{y}})\right] \exp\left[-\frac{1}{2\sigma_\omega^2} \boldsymbol{\omega}^\top \boldsymbol{\omega}\right] \mathbb{I}_{\mathbb{R}_+^N}(\boldsymbol{\omega}) \\
&\propto \exp\Bigg[-\frac{1}{2} \sum_t (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t + \boldsymbol{\omega})^\top [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} \\
&\times (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t + \boldsymbol{\omega})\Bigg] \exp\left[-\frac{1}{2\sigma_\omega^2} \boldsymbol{\omega}^\top \boldsymbol{\omega}\right] \mathbb{I}_{\mathbb{R}_+^N}(\boldsymbol{\omega})
\end{aligned}
$$

$$
\begin{aligned}
&\propto \exp\Bigg[-\frac{1}{2}\Big(\textstyle\sum_t \boldsymbol{\omega}^\top [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} \boldsymbol{\omega} \\
&+ \textstyle\sum_t (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t)^\top [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} \boldsymbol{\omega} \\
&+ \textstyle\sum_t \boldsymbol{\omega}^\top [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t) \Big) \Bigg] \\
&\times \exp\left[-\frac{1}{2\sigma_\omega^2} \boldsymbol{\omega}^\top \boldsymbol{\omega} \right] \mathbb{I}_{\mathbb{R}_+^N}(\boldsymbol{\omega})
\end{aligned}
$$

Define $\mathbf{C}_{\boldsymbol{\omega}}^{-1} = [\kappa_2 \sigma_\varepsilon]^{-1} \sum_t [\mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} + [\sigma_\omega^2 \mathbf{I}_N]^{-1}$. Note that $\mathbf{C}_{\boldsymbol{\omega}}^{-1}$ is diagonal matrix, where the diagonal elements have the form $\left[(\kappa_2 \sigma_\epsilon)^{-1} \sum_t \nu_{it}^{-1} + (\sigma_\omega^2)^{-1} \right]$, for $i = 1, \ldots, \mathrm{N}$. Thus

$$
\begin{aligned}
\pi(\boldsymbol{\omega} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\theta}) \;\propto\; \exp\Bigg[-\frac{1}{2} \big(&\boldsymbol{\omega}^\top \mathbf{C}_{\boldsymbol{\omega}}^{-1} \boldsymbol{\omega} + \textstyle\sum_t (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t)^\top [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} \mathbf{C}_{\boldsymbol{\omega}} \mathbf{C}_{\boldsymbol{\omega}}^{-1} \boldsymbol{\omega}) \\
&+ \boldsymbol{\omega}^\top \mathbf{C}_{\boldsymbol{\omega}}^{-1} \textstyle\sum_t \mathbf{C}_{\boldsymbol{\omega}} [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t) \big) \Bigg] \mathbb{I}_{\mathbb{R}_+^N}(\boldsymbol{\omega}).
\end{aligned}
$$

Define $\mathbf{m}_{\boldsymbol{\omega}} = -\mathbf{C}_{\boldsymbol{\omega}} \sum_t [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t)$. Note that $\sum_t [\kappa_2 \sigma_\varepsilon \mathrm{diag}(\boldsymbol{\nu}_t)]^{-1} (\mathbf{y}_t - \mathbf{x}_t^\top \boldsymbol{\beta} - \kappa_1 \boldsymbol{\nu}_t)$ is vector which elements have the form $\left[\sum_t (y_{it} - x_{it}^\top \boldsymbol{\beta} - \kappa_1 \nu_{it}) / (\kappa_2 \sigma_\varepsilon \nu_{it}) \right]$, for $i = 1, \ldots, \mathrm{N}$. Thus

$$
\pi(\boldsymbol{\omega} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\beta}, \sigma_\varepsilon, \sigma_\omega^2) \;\propto\; \exp\left[-\frac{1}{2} \left((\boldsymbol{\omega} - \mathbf{m}_{\boldsymbol{\omega}})^\top \mathbf{C}_{\boldsymbol{\omega}}^{-1} (\boldsymbol{\omega} - \mathbf{m}_{\boldsymbol{\omega}}) \right) \right] \mathbb{I}_{\mathbb{R}_+^N}(\boldsymbol{\omega}),
$$

which is the kernel of the truncated Gaussian distribution on $\mathbb{R}_+^N$, i.e. $(\boldsymbol{\omega} \mid \mathbf{y}, \boldsymbol{\nu}, \boldsymbol{\beta}, \sigma_\varepsilon, \sigma_\omega) \sim \mathcal{N}_{\mathbb{R}_+^N}(\mathbf{m}_{\boldsymbol{\omega}}, \mathbf{C}_{\boldsymbol{\omega}})$.