Computational Statistics and Data Analysis 54 (2010) 2736–2752

Contents lists available at ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Tobit model with covariate dependent thresholds

Yasuhiro Omori[a,*], Koji Miyawaki[b]

[a] *Faculty of Economics, University of Tokyo, Tokyo 113-0033, Japan*

[b] *Graduate School of Economics, University of Tokyo, Tokyo 113-0033, Japan*

ARTICLE INFO

Article history:
Available online 12 February 2009

ABSTRACT

Tobit models are extended to allow threshold values which depend on individuals' characteristics. In such models, the parameters are subject to as many inequality constraints as the number of observations, and the maximum likelihood estimation which requires the numerical maximisation of the likelihood is often difficult to be implemented. Using a Bayesian approach, a Gibbs sampler algorithm is proposed and, further, the convergence to the posterior distribution is accelerated by introducing an additional scale transformation step. The procedure is illustrated using the simulated data, wage data and prime rate changes' data.

1. Introduction

A censored regression model has been very popular and well known as a standard Tobit (Type I Tobit) model in economics since it was first introduced by Tobin (1958) to analyze the relationship between the household income and household expenditures on a durable good where there are some households with zero expenditures (see e.g., Amemiya (1984) for a survey). The standard Type I Tobit model is given by

$$y_i = \begin{cases} y_i^*, & \text{if } y_i^* \geq d, \\ \text{n.a.,} & \text{otherwise,} \end{cases} \quad i = 1, 2, \ldots, n, \tag{1}$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\alpha} + \epsilon_i, \quad \epsilon_i \sim \text{ i.i.d. } \mathcal{N}(0, \tau^2), \tag{2}$$

where y_i is a dependent variable, y_i^* is a latent dependent variable, d is a censoring limit, $\boldsymbol{x}_i$ is a $K \times 1$ covariate vector and $\boldsymbol{\alpha}$ is a corresponding $K \times 1$ regression coefficient vector. We observe a response variable y_i when it is greater than or equal to a threshold d. The threshold is assumed to be a known constant, and often set equal to zero for convenience. A Bayesian estimation method of such a Tobit model was first proposed by Chib (1992). Chib (1992) developed a Gibbs sampling procedure using the idea of the data augmentation, which is widely used in the literature, and compared the efficacy of the different Monte Carlo methods.

When d is unknown, it will be absorbed into the constant term of the regression, and other coefficients are estimated properly. However, as discussed in Zuehlke (2003), the individual threshold values may be of great interest and need to be estimated separately. For ordered response categories, heterogeneous thresholds' ordinal regression models are considered to account for individual differences in the response style in Johnson (2003) to analyze job preferences and in De Jong et al. (2007) to model consumers' susceptibility to normative influence using a hierarchical item response theory model. In political science, King et al. (2004) also discussed the ordered probit model with thresholds varying over individuals as a function of measured explanatory variables in survey research.

* Corresponding author. Tel.: +81 3 5841 5516; fax: +81 3 5841 5521.
E-mail address: omori@e.u-tokyo.ac.jp (Y. Omori).

0167-9473/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2009.02.005

Thus it is more natural to extend the standard Tobit model such that the deterministic thresholds can vary with individuals depending on their characteristics. In such a model with covariate dependent thresholds, the ith response variable y_i is observed if it is greater than or equal to a threshold $d_i = \boldsymbol{w}_i'\boldsymbol{\delta}$ where $\boldsymbol{w}_i$ and $\boldsymbol{\delta}$ are a $J \times 1$ covariate vector and a corresponding coefficient vector, respectively.

The alternative extension of the standard Tobit model is known as a sample selection model or a generalised Tobit (Type II Tobit) model in the literature,

$$y_i = \begin{cases} y_i^*, & \text{if } z_i^* \geq 0, \\ \text{n.a.}, & \text{otherwise}, \end{cases} \quad i = 1, 2, \ldots, n, \tag{3}$$

$$z_i^* = \boldsymbol{w}_i'\boldsymbol{\theta} + \xi_i, \tag{4}$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\beta} + \eta_i, \tag{5}$$

where $(\xi_i, \eta_i)' \sim$ i.i.d. $\mathcal{N}(0, \Sigma)$, $(\boldsymbol{w}_i, \boldsymbol{x}_i)$ are independent variable vectors, $(\boldsymbol{\theta}, \boldsymbol{\beta})$ are corresponding coefficient vectors and the $(1, 1)$ element of Σ is set equal to 1 for the identification. The sample rule is determined by a latent random variable z_i^*, and we observe the response variable y_i when $z_i^* \geq 0$. The latent variable z_i^* is allowed to be correlated with the response variable y_i^*. When the correlation coefficient, ρ, between (z_i^*, y_i^*) is not equal to zero, a sample selection model is considered a Tobit model with a stochastic threshold model.

However, the correlation coefficient is often estimated to be almost one in the empirical studies (see e.g., Table 7 in Section 5.1). If ρ is equal to one, the generalised Tobit model reduces to the special case of the Tobit model with covariate dependent thresholds which we just mentioned above. Whether we should use a standard Tobit model, a sample selection model or a Tobit model with covariate dependent thresholds has been an important issue in the empirical studies (see Cragg (1971), Lin and Schmidt (1984) and Melenberg and Soest (1996)). In this paper, we take a Bayesian approach to deal with such a problem using the DIC (Deviance Information Criterion, see e.g., Spiegelhalter et al. (2002), Congdon (2005) and Shriner and Yi (2008)) as a model selection criterion.

The purpose of this paper is two-fold. First we propose a Markov chain Monte Carlo (MCMC) estimation method for a Tobit model with the thresholds which depend on individuals' characteristics. Since the parameters are subject to as many inequality constraints as the number of observations, the numerical maximisation of the likelihood is often difficult to be implemented. Using a Bayesian approach, we describe a Gibbs sampler algorithm to estimate parameters, and, further, show that the speed of convergence to the posterior distribution can be accelerated by adding one more step to the simple Gibbs sampler.

Second, we extend our proposed model to a friction model introduced by Rosett (1959) (see also Maddala (1983)) with covariate dependent thresholds, and a two-limit Tobit model with covariate dependent thresholds. The friction model has been used, for example, to analyze the nominal and real wage rigidity (Christofidesc and Li, 2005), the prime rate change (Forbes and Mayne, 1989), and the price change of a product or a brand over time (Desarbo et al., 1987). The dependent variable is subject to friction, and is not affected by the independent variables in the regression if the changes in the regression equation are small. It is affected only when the changes in the regression equation are greater than the upper limit or smaller than the lower limit. These limits are usually assumed to be constants, but we will consider the covariate dependent limits to extend the standard friction model.

Our proposed estimation methods are illustrated using both simulated data and real data. The real data examples include the popular example of hourly wage of married women (discussed in Mroz (1987)) and prime rate changes in Japan. The model comparison is conducted using the DIC for wage data, while the posterior predictive checks are used to assess the plausibility of our proposed friction model for prime rate changes' data.

The rest of the paper is organised as follows. In Section 2, we propose an efficient Gibbs sampler for a Tobit model with covariate dependent thresholds and, in Section 3, extensions to a friction model and a two-limit Tobit model are discussed. Section 4 illustrates our proposed estimation method using simulated data. Empirical studies are shown in Section 5 using wage data and prime rate changes' data. Section 6 concludes the paper.

2. Tobit model with covariate dependent thresholds

2.1. Gibbs sampler

We first describe a Gibbs sampler for a Tobit (standard Tobit Type 1) model (see e.g., Chib (1995)). The prior distributions of $(\boldsymbol{\alpha}, \tau^2)$ are assumed to be a conditionally multivariate normal distribution and an inverse gamma distribution, respectively,

$$\boldsymbol{\alpha}|\tau^2 \sim \mathcal{N}(\boldsymbol{a}_0, \tau^2 A_0), \quad \tau^2 \sim \mathcal{IG}\left(\frac{n_0}{2}, \frac{S_0}{2}\right), \tag{6}$$

where $\boldsymbol{a}_0$ is a $K \times 1$ known constant vector, A_0 is a $K \times K$ known constant matrix, and n_0, S_0 are known positive constants. To implement a Markov chain Monte Carlo method, we use a data augmentation method by sampling an unobserved latent response variable y_i^*. Using $\boldsymbol{y}^*$, model (1)–(2) reduces to an ordinary linear regression model, $\boldsymbol{y}^* = X\boldsymbol{\alpha} + \boldsymbol{\epsilon}$, where

$\boldsymbol{y}^* = (y_1^*, y_2^*, \dots, y_n^*)'$, $X' = (\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n)$ and $\boldsymbol{\epsilon} = (\epsilon_1, \epsilon_2, \dots, \epsilon_n)' \sim \mathcal{N}(0, \tau^2 I_n)$. Given $\boldsymbol{y}^*$, the conditional posterior distributions of $(\boldsymbol{\alpha}, \tau^2)$ are

$$\boldsymbol{\alpha}|\tau^2, \boldsymbol{y}^* \sim \mathcal{N}(\boldsymbol{a}_1, \tau^2 A_1), \quad \tau^2|\boldsymbol{y}^* \sim \mathcal{IG}\left(\frac{n_1}{2}, \frac{S_1}{2}\right),$$

where $A_1^{-1} = A_0^{-1} + X'X$, $\boldsymbol{a}_1 = A_1(A_0^{-1}\boldsymbol{a}_0 + X'\boldsymbol{y}^*)$, $n_1 = n_0 + n$, and $S_1 = \boldsymbol{y}^{*'}\boldsymbol{y}^* + \boldsymbol{a}_0' A_0^{-1}\boldsymbol{a}_0 + S_0 - \boldsymbol{a}_1' A_1^{-1}\boldsymbol{a}_1$ (see Appendix A.1). Let $\boldsymbol{y}_o = (y_{o,1}, y_{o,2}, \dots, y_{o,m})'$ and $\boldsymbol{y}_c^* = (y_{c,1}^*, y_{c,2}^*, \dots, y_{c,n-m}^*)'$ denote $m \times 1$ and $(n-m) \times 1$ vectors of observed (uncensored) and censored dependent variables, respectively. Then, we can sample from the posterior distribution using a Gibbs sampler in two blocks:

(1) Initialise $\boldsymbol{\alpha}$ and τ^2.
(2) Sample $\boldsymbol{y}_c^*|\boldsymbol{\alpha}, \tau^2, \boldsymbol{y}_o$.
    Generate $y_{c,i}^*|\boldsymbol{\alpha}, \tau^2 \sim \mathcal{TN}_{(-\infty,d)}(\boldsymbol{x}_i'\boldsymbol{\alpha}, \tau^2)$, $i = 1, 2, \dots, n-m$, for censored observations, where $\mathcal{TN}_{(a,b)}(\mu, \sigma^2)$ denotes a normal distribution $\mathcal{N}(\mu, \sigma^2)$ truncated on the interval (a, b).
(3) Sample $(\boldsymbol{\alpha}, \tau^2)|\boldsymbol{y}_c^*, \boldsymbol{y}_o$
    (a) Sample $\tau^2|\boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{IG}(n_1/2, S_1/2)$,
    (b) Sample $\boldsymbol{\alpha}|\tau^2, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{N}(\boldsymbol{a}_1, \tau^2 A_1)$.
(4) Go to 2.

Next, we extend it to the model with covariate dependent thresholds model. By adding another block to the above sampler, we can derive the Gibbs sampler for the Tobit model with covariate dependent thresholds. In the standard Tobit model (1)–(2), the threshold d is assumed to be known and a constant. However, it is usually unknown and may vary with the individual characteristics. Thus we extend it to allow unknown but covariate dependent thresholds as follows.

$$y_i = \begin{cases} y_i^*, & \text{if } y_i^* \ge \boldsymbol{w}_i'\boldsymbol{\delta}, \\ \text{n.a.}, & \text{otherwise}, \end{cases} \quad i = 1, 2, \dots, n, \tag{7}$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\alpha} + \epsilon_i, \quad \epsilon_i \sim \text{ i.i.d. } \mathcal{N}(0, \tau^2), \tag{8}$$

where $(\boldsymbol{w}_i, \boldsymbol{x}_i)$ are $J \times 1$ and $K \times 1$ covariate vectors and $(\boldsymbol{\delta}, \boldsymbol{\alpha})$ are corresponding $J \times 1$ and $K \times 1$ regression coefficient vectors. The known constant threshold d in (1) is replaced by the unknown but covariate dependent threshold, $\boldsymbol{w}_i'\boldsymbol{\delta}$.

To conduct a Bayesian analysis of the proposed Tobit model (7)–(8), we assume that prior distributions of $(\boldsymbol{\alpha}, \tau^2)$ are given by (6). A prior distribution of $\boldsymbol{\delta}$ is assumed to be $\boldsymbol{\delta}|\tau^2 \sim \mathcal{N}(\boldsymbol{d}_0, \tau^2 D_0)$ since we often use similar independent variables for $\boldsymbol{w}_i$'s to those for $\boldsymbol{x}_i$'s (or standardized independent variables as in our empirical studies in Section 5), and the magnitude of the dispersion is expected to be similar. If there is little prior information with respect to $\boldsymbol{\delta}$, we take large values for the diagonal elements of D_0, which will result in a fairly flat prior for $\boldsymbol{\delta}$.

Then, the conditional posterior distributions of $(\boldsymbol{\alpha}, \tau^2, \boldsymbol{\delta})$ are

$$\boldsymbol{\alpha}|\boldsymbol{\delta}, \tau^2, \boldsymbol{y}^* \sim \mathcal{N}(\boldsymbol{a}_1, \tau^2 A_1), \quad \tau^2|\boldsymbol{\delta}, \boldsymbol{\alpha}, \boldsymbol{y}^* \sim \mathcal{IG}\left(\frac{n_1}{2}, \frac{S_1}{2}\right),$$

$$\boldsymbol{\delta}|\boldsymbol{\alpha}, \tau^2, \boldsymbol{y}^* \sim \mathcal{TN}_{R_o \cap R_c}(\boldsymbol{d}_0, \tau^2 D_0),$$

where $n_1 = n_0 + n + J$, $S_1 = \boldsymbol{y}^{*'}\boldsymbol{y}^* + \boldsymbol{a}_0' A_0^{-1}\boldsymbol{a}_0 - \boldsymbol{a}_1' A_1^{-1}\boldsymbol{a}_1 + S_0 + (\boldsymbol{\delta} - \boldsymbol{d}_0)' D_0^{-1}(\boldsymbol{\delta} - \boldsymbol{d}_0)$, $A_1^{-1} = A_0^{-1} + X'X$, $\boldsymbol{a}_1 = A_1(A_0^{-1}\boldsymbol{a}_0 + X'\boldsymbol{y}^*)$, $R_o = \{\boldsymbol{\delta}|\boldsymbol{w}_i'\boldsymbol{\delta} \le y_i \text{ for uncensored } i\}$, $R_c = \{\boldsymbol{\delta}|\boldsymbol{w}_i'\boldsymbol{\delta} > y_i^* \text{ for censored } i\}$ (see Appendix A.2). The Gibbs sampler is implemented in three blocks as follows.

(1) Initialise $\boldsymbol{\delta}$, $\boldsymbol{\alpha}$ and τ^2 where $\boldsymbol{\delta} \in R_o$.
(2) Sample $\boldsymbol{y}_c^*|\boldsymbol{\alpha}, \tau^2, \boldsymbol{\delta}, \boldsymbol{y}_o$. Generate $y_{c,i}^*|\boldsymbol{\delta}, \boldsymbol{\alpha}, \tau^2 \sim \mathcal{TN}_{(-\infty, \boldsymbol{w}_i'\boldsymbol{\delta})}(\boldsymbol{x}_i'\boldsymbol{\alpha}, \tau^2)$, $i = 1, 2, \dots, n-m$, for censored observations.
(3) Sample $(\boldsymbol{\alpha}, \tau^2)|\boldsymbol{\delta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o$
    (a) Sample $\tau^2|\boldsymbol{\delta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{IG}(n_1/2, S_1/2)$,
    (b) Sample $\boldsymbol{\alpha}|\tau^2, \boldsymbol{\delta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{N}(\boldsymbol{a}_1, \tau^2 A_1)$.
(4) Sample $\boldsymbol{\delta}|\boldsymbol{\alpha}, \tau^2, \boldsymbol{y}^* \sim \mathcal{TN}_{R_o \cap R_c}(\boldsymbol{d}_0, \tau^2 D_0)$.
(5) Go to 2.

Steps 2 and 3 are similar to those in the simple Tobit model. To sample from the conditional posterior distribution of $\boldsymbol{\delta}$ in Step 4, we generate one component δ_j of $\boldsymbol{\delta} = (\delta_1, \delta_2, \dots, \delta_J)'$ at a time given other components $\boldsymbol{\delta}_{-j} = (\delta_1, \dots, \delta_{j-1}, \delta_{j+1} \dots, \delta_J)'$. Since $\boldsymbol{\delta}$ should lie in the region $R_o \cap R_c$, the δ_j is subject to the constraint $L_j \le \delta_j \le U_j$ where $\boldsymbol{w}_{i,-j} = (w_{i1}, \dots, w_{i,j-1}, w_{i,j+1}, \dots, w_{iJ})'$,

$$L_j = \max_i L_{ij}, \quad L_{ij} = \begin{cases} w_{ij}^{-1}(y_i - \boldsymbol{w}_{i,-j}'\boldsymbol{\delta}_{-j}) & \text{if } w_{ij} < 0 \text{ for uncensored } i, \\ w_{ij}^{-1}(y_i^* - \boldsymbol{w}_{i,-j}'\boldsymbol{\delta}_{-j}) & \text{if } w_{ij} > 0 \text{ for censored } i, \\ -\infty, & \text{otherwise}, \end{cases}$$

$$U_j = \min_i U_{ij}, \quad U_{ij} = \begin{cases} w_{ij}^{-1}(y_i - \boldsymbol{w}_{i,-j}'\boldsymbol{\delta}_{-j}) & \text{if } w_{ij} > 0 \text{ for uncensored } i, \\ w_{ij}^{-1}(y_i^* - \boldsymbol{w}_{i,-j}'\boldsymbol{\delta}_{-j}) & \text{if } w_{ij} < 0 \text{ for censored } i, \\ +\infty, & \text{otherwise}. \end{cases}$$

Let $\boldsymbol{d}_{0,-j} = (d_{01}, \ldots, d_{0,j-1}, d_{0,j+1}, \ldots, d_{0J})'$ and let $D_{0,j,j}$, $D_{0,j,-j}$, and $D_{0,-j,-j}$ denote a prior variance of δ_j, a prior covariance vector of δ_j and $\boldsymbol{\delta}_{-j}$, and a prior covariance matrix of $\boldsymbol{\delta}_{-j}$, respectively. Then, we sample δ_j, for $j = 1, 2, \ldots, J$, using the conditional truncated normal posterior distribution,

$$\delta_j|\boldsymbol{\delta}_{-j}, \boldsymbol{\alpha}, \tau^2, \boldsymbol{y}^* \sim \mathcal{TN}_{(L_j,U_j)}(m_j, s_j^2\tau^2),$$
$$m_j = d_{0j} + D_{0,j,-j}D_{0,-j,-j}^{-1}(\boldsymbol{\delta}_{-j} - \boldsymbol{d}_{0,-j}),$$
$$s_j^2 = D_{0,j,j} - D_{0,j,-j}D_{0,-j,-j}^{-1}D_{0,j,-j}'.$$

Note that this reduces to $\mathcal{TN}_{(L_j,U_j)}(d_{0j}, \tau^2 D_{0,j,j})$ for a diagonal D_0.

2.2. Acceleration of the Gibbs sampler

The above sampling scheme for $\boldsymbol{\delta}$ is to sample one component at a time which is known to be an inefficient method, and the obtained samples may exhibit high autocorrelations (see Sections 4.1 and 5.1).

To reduce such high sample autocorrelations, we consider a generalised Gibbs move discussed in Liu and Sabatti (2000). Using a generalised Gibbs move, we can accelerate the convergence of the distribution of the MCMC samples by transforming some parameters and latent variables simultaneously without changing the stationary distribution of the Markov chain. The possible transformation groups are, for example, the scale group, the affine transformation group and the orthonormal transformation group. Since parameters and latent variables are subject to many inequality constraints in our model, we consider the scale group so that most inequality constraints remain satisfied for those transformed parameters and variables. With the scale group, we would have less constraints on the auxiliary variable which we use for the transformation.

Consider the scale group $\Gamma = \{g > 0 : g(\boldsymbol{\delta}, \boldsymbol{\alpha}, \tau, \boldsymbol{y}_c^*) = (g\boldsymbol{\delta}, g\boldsymbol{\alpha}, g\tau, g\boldsymbol{y}_c^*)\}$ (see Appendix B), and let $\boldsymbol{x}_{o,i}$ denote the independent variable vector of the ith observed response $y_{o,i}$ and $X_o' = (\boldsymbol{x}_{o,1}, \ldots, \boldsymbol{x}_{o,m})$. Noting that $\lambda_L < g^{-1} < \lambda_U$ where

$$\lambda_L = \max_i L_{\lambda,i}, \quad L_{\lambda,i} = \begin{cases} \max\left(0, \dfrac{\boldsymbol{w}_{o,i}'\boldsymbol{\delta}}{y_{o,i}}\right), & \text{if } y_{o,i} > 0, \\ 0, & \text{otherwise,} \end{cases}$$
$$\lambda_U = \min_i U_{\lambda,i}, \quad U_{\lambda,i} = \begin{cases} \dfrac{\boldsymbol{w}_{o,i}'\boldsymbol{\delta}}{y_{o,i}}, & \text{if } y_{o,i} < 0, \\ \infty, & \text{otherwise,} \end{cases}$$

the acceleration steps are given as follows. When $\lambda_U < \infty$:

(5) (a) Generate $g^{-1} \sim \mathcal{U}(\lambda_L, \lambda_U)$ where $\mathcal{U}(a, b)$ denotes a uniform distribution on the interval (a, b).
   (b) Accept g with probability
$$\min\left[1, g^{-(n_0+m-1)} \exp\left\{-\frac{-2q(g^{-1}-1) + r(g^{-2}-1)}{2\tau^2}\right\}\right],$$
   where $q = \boldsymbol{\alpha}'X_o'\boldsymbol{y}_o + \boldsymbol{\alpha}'A_0^{-1}\boldsymbol{a}_0 + \boldsymbol{d}_0'D_0^{-1}\boldsymbol{\delta}$, $r = \boldsymbol{y}_o'\boldsymbol{y}_o + \boldsymbol{a}_0'A_0^{-1}\boldsymbol{a}_0 + \boldsymbol{d}_0'D_0^{-1}\boldsymbol{d}_0 + S_0$. If rejected, set $g = 1$.
   (c) Let $g\boldsymbol{\delta} \to \boldsymbol{\delta}$, $g\boldsymbol{\alpha} \to \boldsymbol{\alpha}$, $g\tau \to \tau$ and $g\boldsymbol{y}_c^* \to \boldsymbol{y}_c^*$.

(6) Go to 2.

When $\lambda_U = \infty$, we replace Step 5 (a) (b) by:

(5) (a)′ Generate $g^{-1} \sim \mathcal{TN}_{(\lambda_L,\infty)}(1, \sigma_\lambda^2)$ where σ_λ^2 is some specified constant (e.g., $\sigma_\lambda^2 = 0.5^2$).
   (b)′ Accept g with probability
$$\min\left[1, g^{-(n_0+m-1)} \exp\left\{-\frac{-2q(g^{-1}-1) + r(g^{-2}-1)}{2\tau^2} + \frac{(g^{-1}-1)^2}{2\sigma_\lambda^2}\right\}\right].$$
   If rejected, set $g = 1$.

We illustrate how effective this acceleration step is to improve the speed of the convergence (of the distribution of the MCMC samples) to the target posterior distribution in Sections 4.1 and 5.1.

3. Extensions

In this section, we describe two useful econometric models which can be obtained as extensions of our basic covariate dependent thresholds model (7) and (8).

3.1. Friction model

The first extension is for a friction model introduced by Rosett (1959). The friction model with covariate dependent thresholds is given by

$$y_i = \begin{cases} y_i^* - \boldsymbol{w}_i'\boldsymbol{\delta}, & \text{if } y_i^* - \boldsymbol{w}_i'\boldsymbol{\delta} < c, \\ c, & \text{if } \boldsymbol{w}_i'\boldsymbol{\delta} + c \le y_i^* \le \boldsymbol{v}_i'\boldsymbol{\zeta} + c, \\ y_i^* - \boldsymbol{v}_i'\boldsymbol{\zeta}, & \text{if } y_i^* - \boldsymbol{v}_i'\boldsymbol{\zeta} > c, \end{cases} \quad i = 1, 2, \ldots, n, \tag{9}$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\alpha} + \epsilon_i, \quad \epsilon_i \sim \text{i.i.d. } \mathcal{N}(0, \tau^2). \tag{10}$$

where c is a known constant, and $(\boldsymbol{v}_i, \boldsymbol{\zeta})$ are $L \times 1$ covariate and corresponding coefficient vectors. Since c is set equal to 0 in most applications, we focus on the friction model with $c = 0$. Let R_δ and R_ζ denote truncation regions defined by

$$R_\delta = R_{\delta,o} \cap R_{\delta,c}, \quad R_\zeta = R_{\zeta,o} \cap R_{\zeta,c},$$
$$R_{\delta,o} = \left\{\boldsymbol{\delta} \mid \boldsymbol{w}_i'\boldsymbol{\delta} \le \boldsymbol{v}_i'\boldsymbol{\zeta} \text{ for uncensored } i\right\}, \quad R_{\zeta,o} = \left\{\boldsymbol{\zeta} \mid \boldsymbol{w}_i'\boldsymbol{\delta} \le \boldsymbol{v}_i'\boldsymbol{\zeta} \text{ for uncensored } i\right\},$$
$$R_{\delta,c} = \left\{\boldsymbol{\delta} \mid \boldsymbol{w}_i'\boldsymbol{\delta} \le y_i^* \text{ for censored } i\right\}, \quad R_{\zeta,c} = \left\{\boldsymbol{\zeta} \mid y_i^* \le \boldsymbol{v}_i'\boldsymbol{\zeta} \text{ for censored } i\right\}.$$

Assuming that the conditionally normal prior distribution for $\boldsymbol{\zeta}$, $\boldsymbol{\zeta} \mid \tau^2 \sim \mathcal{N}(\boldsymbol{z}_0, \tau^2 Z_0)$ and the same prior distributions for other parameters are as in the previous section (a conditionally normal prior for $\boldsymbol{\alpha}$ and $\boldsymbol{\delta}$, an inverse gamma prior for τ^2), we implement the Gibbs sampler in six blocks:

(1) Initialise $\boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{\alpha}$ and τ^2 where $\boldsymbol{\delta} \in R_{\delta,o}$ and $\boldsymbol{\zeta} \in R_{\zeta,o}$.
(2) Sample $\boldsymbol{y}_c^* | \boldsymbol{\alpha}, \tau^2, \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_o$. Generate $y_{c,i}^* | \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{\alpha}, \tau^2 \sim \mathcal{TN}_{[\boldsymbol{w}_i'\boldsymbol{\delta}, \boldsymbol{v}_i'\boldsymbol{\zeta}]}(\boldsymbol{x}_i'\boldsymbol{\alpha}, \tau^2)$, $i = 1, 2, \ldots, n - m$, for censored observations.
(3) Sample $(\boldsymbol{\alpha}, \tau^2) | \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o$
 (a) Sample $\tau^2 | \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{IG}(n_1/2, S_1/2)$ where $n_1 = n_0 + J + L + n$ and $S_1 = \boldsymbol{y}^{*\prime}\boldsymbol{y}^* + \boldsymbol{a}_0'\mathbf{A}_0^{-1}\boldsymbol{a}_0 - \boldsymbol{a}_1'\mathbf{A}_1^{-1}\boldsymbol{a}_1 + S_0 + (\boldsymbol{\delta} - \boldsymbol{d}_0)'\mathbf{D}_0^{-1}(\boldsymbol{\delta} - \boldsymbol{d}_0) + (\boldsymbol{\zeta} - \boldsymbol{z}_0)'\mathbf{Z}_0^{-1}(\boldsymbol{\zeta} - \boldsymbol{z}_0)$.
 (b) Sample $\boldsymbol{\alpha} | \tau^2, \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{N}(\boldsymbol{a}_1, \tau^2 A_1)$ where $\mathbf{A}_1^{-1} = \mathbf{A}_0^{-1} + \mathbf{X}'\mathbf{X}$ and $\boldsymbol{a}_1 = \mathbf{A}_1(\mathbf{A}_0^{-1}\boldsymbol{a}_0 + \mathbf{X}'\boldsymbol{y}^*)$.
(4) Sample $\boldsymbol{\delta} | \boldsymbol{\zeta}, \boldsymbol{\alpha}, \tau^2, \boldsymbol{y}^* \sim \mathcal{TN}_{R_\delta}(\boldsymbol{d}_1, \tau^2 D_1)$ where $\mathbf{D}_1^{-1} = \mathbf{D}_0^{-1} + \sum_{i \in \mathcal{O}^-} \boldsymbol{w}_i\boldsymbol{w}_i'$, $\boldsymbol{d}_1 = \mathbf{D}_1\left\{\mathbf{D}_0^{-1}\boldsymbol{d}_0 + \sum_{i \in \mathcal{O}^-}(\boldsymbol{x}_i'\boldsymbol{\alpha} - y_i)\boldsymbol{w}_i\right\}$ and $\mathcal{O}^- = \{i \mid y_i < 0\}$.
(5) Sample $\boldsymbol{\zeta} | \boldsymbol{\delta}, \boldsymbol{\alpha}, \tau^2, \boldsymbol{y}^* \sim \mathcal{TN}_{R_\zeta}(\boldsymbol{z}_1, \tau^2 Z_1)$ where $\mathbf{Z}_1^{-1} = \mathbf{Z}_0^{-1} + \sum_{i \in \mathcal{O}^+} \boldsymbol{v}_i\boldsymbol{v}_i'$, $\boldsymbol{z}_1 = \mathbf{Z}_1\left\{\mathbf{Z}_0^{-1}\boldsymbol{z}_0 + \sum_{i \in \mathcal{O}^+}(\boldsymbol{x}_i'\boldsymbol{\alpha} - y_i)\boldsymbol{v}_i\right\}$ and $\mathcal{O}^+ = \{i \mid y_i > 0\}$.
(6) Go to 2.

3.2. Two-limit Tobit model

The second extension is for a two-limit Tobit model in which the dependent variable is doubly censored with known limits (see e.g., Maddala (1983)). The model has been applied in the various empirical studies, such as the estimation of the electric vehicle demand share equation (Hill, 1987), and the stock price exchange movements with limits (Charemza and Majerowska, 2000). The extended two-limit model is given by

$$y_i = \begin{cases} y_i^*, & \text{if } \boldsymbol{w}_i'\boldsymbol{\delta} \le y_i^* \le \boldsymbol{v}_i'\boldsymbol{\zeta}, \\ \text{n.a.}, & \text{otherwise}, \end{cases} \quad i = 1, 2, \ldots, n, \tag{11}$$

$$y_i^* = \boldsymbol{x}_i'\boldsymbol{\alpha} + \epsilon_i, \quad \epsilon_i \sim \text{i.i.d. } \mathcal{N}(0, \tau^2), \tag{12}$$

where y_i is observed only when it falls within the interval, and the upper and lower limits depend on the individual's characteristics.

Assuming the same prior distributions as those in the previous friction model (a conditionally normal prior for $\boldsymbol{\alpha}, \boldsymbol{\delta}, \boldsymbol{\zeta}$, an inverse gamma prior for τ^2), we obtain the Gibbs sampling algorithm as follows.

(1) Initialise $\boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{\alpha}$ and τ^2 where $\boldsymbol{\delta} \in R_{\delta,o}$ and $\boldsymbol{\zeta} \in R_{\zeta,o}$.
(2) Sample $\boldsymbol{y}_c^* | \boldsymbol{\alpha}, \tau^2, \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_o$. Generate $y_{c,i}^* | \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{\alpha}, \tau^2 \sim \mathcal{TN}_{(-\infty, \boldsymbol{w}_i'\boldsymbol{\delta}) \cup (\boldsymbol{v}_i'\boldsymbol{\zeta}, \infty)}(\boldsymbol{x}_i'\boldsymbol{\alpha}, \tau^2)$, $i = 1, 2, \ldots, n - m$, for censored observations.
(3) Sample $(\boldsymbol{\alpha}, \tau^2) | \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o$
 (a) Sample $\tau^2 | \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{IG}(n_1/2, S_1/2)$,
 (b) Sample $\boldsymbol{\alpha} | \tau^2, \boldsymbol{\delta}, \boldsymbol{\zeta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{N}(\boldsymbol{a}_1, \tau^2 A_1)$.
(4) Sample $\boldsymbol{\delta} | \boldsymbol{\zeta}, \boldsymbol{\alpha}, \tau^2, \boldsymbol{y}^* \sim \mathcal{TN}_{R_\delta}(\boldsymbol{d}_0, \tau^2 D_0)$.
(5) Sample $\boldsymbol{\zeta} | \boldsymbol{\delta}, \boldsymbol{\alpha}, \tau^2, \boldsymbol{y}^* \sim \mathcal{TN}_{R_\zeta}(\boldsymbol{z}_0, \tau^2 Z_0)$.
(6) Go to 2.

Table 1
Posterior means, standard deviations, 95% credible intervals and inefficiency factors obtained from simple Gibbs sampler.

	True	Mean	Stdev	95% Interval	Inef
δ_1	1.0	0.968	0.062	(0.868, 1.088)	6.6
δ_2	5.0	4.933	0.204	(4.641, 5.448)	219.8
δ_3	10.0	9.758	0.448	(9.124, 10.904)	219.5
α_1	2.0	2.022	0.049	(1.926, 2.119)	1.6
α_2	1.0	0.974	0.051	(0.874, 1.073)	3.3
α_3	1.0	1.066	0.048	(0.972, 1.161)	3.5
τ^2	0.6	0.624	0.058	(0.520, 0.750)	3.4

Table 2
Posterior means, standard deviations, 95% credible intervals and inefficiency factors obtained from accelerated Gibbs sampler.

	True	Mean	Stdev	95% Interval	Inef
δ_1	1.0	0.968	0.063	(0.868, 1.089)	4.9
δ_2	5.0	4.928	0.191	(4.619, 5.363)	44.1
δ_3	10.0	9.747	0.418	(9.080, 10.699)	46.2
α_1	2.0	2.021	0.050	(1.923, 2.118)	2.4
α_2	1.0	0.974	0.050	(0.879, 1.073)	1.7
α_3	1.0	1.065	0.048	(0.971, 1.160)	1.9
τ^2	0.6	0.625	0.058	(0.521, 0.746)	3.9

where $n_1, S_1, \boldsymbol{a}_1, \mathbf{A}_1$ are the same as those for the friction model (see Section 3.1), and

$$\begin{aligned}
&R_\delta = R_{\delta,o} \cap R_{\delta,lc} \cap R_{\delta,rc}, \quad R_\zeta = R_{\zeta,o} \cap R_{\zeta,lc} \cap R_{\zeta,rc},\\
&R_{\delta,o} = \{\boldsymbol{\delta} |\ \boldsymbol{w}_i'\boldsymbol{\delta} \le y_i \text{ for uncensored } i\}, \quad R_{\delta,lc} = \{\boldsymbol{\delta} |\ \boldsymbol{w}_i'\boldsymbol{\delta} > y_i^* \text{ for left censored } i\},\\
&R_{\delta,rc} = \{\boldsymbol{\delta} |\ \boldsymbol{w}_i'\boldsymbol{\delta} < \boldsymbol{v}_i'\boldsymbol{\zeta} \text{ for right censored } i\}, \quad R_{\zeta,o} = \{\boldsymbol{\zeta} |\ \boldsymbol{v}_i'\boldsymbol{\zeta} \ge y_i \text{ for uncensored } i\},\\
&R_{\zeta,lc} = \{\boldsymbol{\zeta} |\ \boldsymbol{w}_i'\boldsymbol{\delta} < \boldsymbol{v}_i'\boldsymbol{\zeta} \text{ for left censored } i\}, \quad R_{\zeta,rc} = \{\boldsymbol{\zeta} |\ \boldsymbol{v}_i'\boldsymbol{\zeta} < y_i^* \text{ for right censored } i\}.
\end{aligned}$$

4. Illustrative examples

In this section, we consider two examples, (i) a Tobit model with covariate dependent thresholds and (ii) an extended friction model, to illustrate our estimation procedure using simulated data.

4.1. Tobit model with covariate dependent thresholds

First, we consider a Tobit model with covariate dependent thresholds given by (7) and (8). Let true values of the model be $\boldsymbol{\delta} = (1, 5, 10)'$, $\boldsymbol{\alpha} = (2, 1, 1)'$, and $\tau^2 = 0.6$. All covariates are generated using a standard normal distribution, *i.e.*, $x_{ij} \sim$ i.i.d. $\mathcal{N}(0, 1)$. The number of generated observations is five hundred, and 53% of them were censored. We assumed the prior distribution for $\boldsymbol{\delta}$, $\boldsymbol{\alpha}$ and τ as follows:

$$\boldsymbol{\delta}|\tau^2 \sim \mathcal{N}(0, 10\tau^2 I_3), \quad \boldsymbol{\alpha}|\tau^2 \sim \mathcal{N}(0, 10\tau^2 I_3), \quad \tau^2 \sim \mathcal{IG}(0.1, 0.1).$$

In the Gibbs sampling from the posterior distribution, the initial 2000 variates are discarded as the burn-in period and the subsequent 10, 000 values are retained.

Table 1 shows the true parameter values, posterior means, standard deviations, 95 % credible intervals, and inefficiency factors for the simple Gibbs sampler. The inefficiency factor is defined as $1+2\sum_{s=1}^{\infty}\rho_s$ where ρ_s is the sample autocorrelation at lag s, and are computed to measure how well the MCMC chain mixes (see e.g., Chib (2001)). It is the ratio of the numerical variance of the posterior sample mean to the variance of the sample mean from uncorrelated draws. The inverse of inefficiency factor is also known as the relative numerical efficiency (Geweke, 1992). When the inefficiency factor is equal to m, we need to draw MCMC samples m times as many as uncorrelated samples.

The posterior means are close to the true values, and all true values are contained in the 95% credible intervals. The inefficiency factors are 1–220 which seem to increase as the value of δ_i becomes large. This is probably because the posterior distribution for large δ_i's becomes more sensitive to the linear inequality constraints.

In Table 2, the corresponding summary statistics for the accelerated Gibbs sampler are shown. The inefficiency factors of δ_i's are 4–47 suggesting that such an acceleration step can improve the mixing property of the Markov chain.

Fig. 1 shows the sample paths, sample autocorrelations functions, and estimated marginal posterior densities from the simple Gibbs sampler. While the sample autocorrelations vanish quickly for $\boldsymbol{\alpha}$ and τ^2, those of $\boldsymbol{\delta}$ do not vanish until 1500 lags indicating the slow convergence of the distribution of the MCMC samples to the posterior distribution.

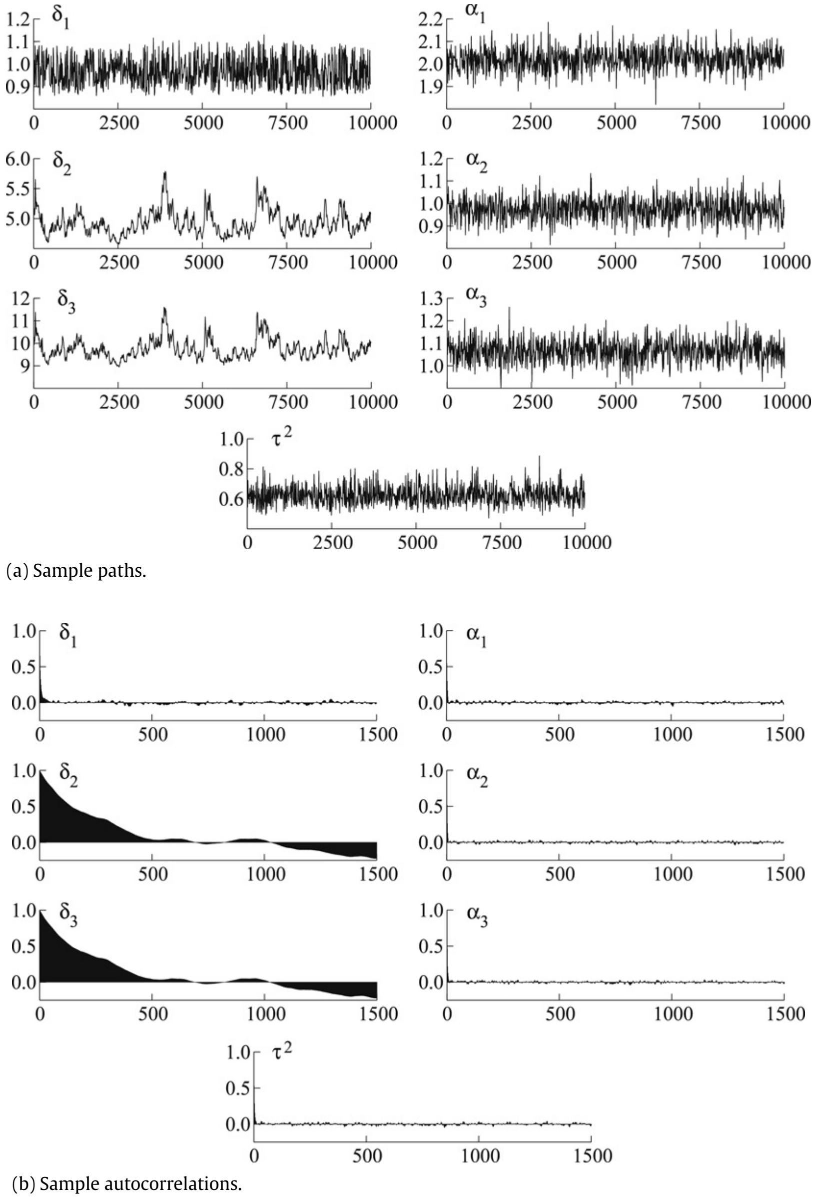


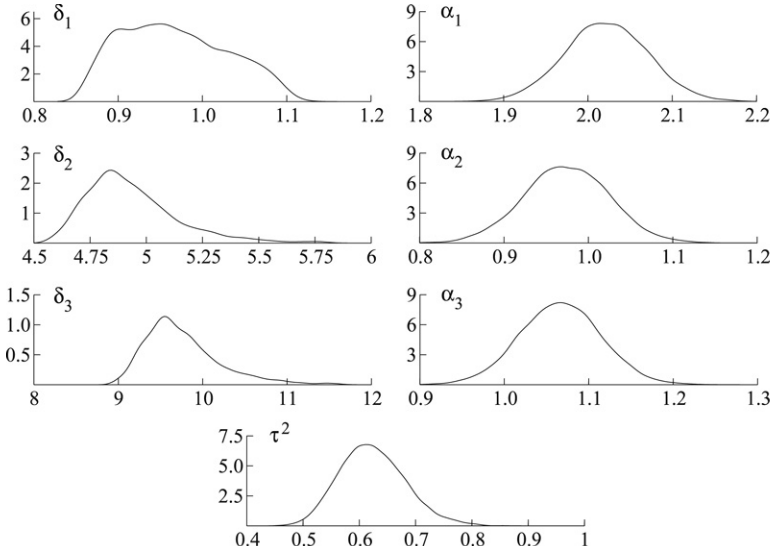
Fig. 1. Tobit model. Simple Gibbs sampler.

On the other hand, as shown in Fig. 2, the sample paths from the accelerated Gibbs sampler indicate that the Markov chains are mixing very well and all sample autocorrelations decay very quickly. It also shows that the acceleration step is effective to improve the speed of the convergence to the target distribution in the MCMC implementation.

4.2. Friction model

Next we illustrate our procedure for the friction model defined by (9) and (10). True values are $\boldsymbol{\delta} = (-1, -0.4)'$, $\boldsymbol{\zeta} = (0.5, 0.2)'$, $\boldsymbol{\alpha} = (2, 0.5)'$, and $\tau^2 = 0.49$. Covariates are generated using a standard normal distribution except for the second covariates of $\boldsymbol{w}_i$ and $\boldsymbol{v}_i$ (*i.e.*, w_{i2} and v_{i2}) where they are obtained by taking the absolute value of standard normal random variables. We generated five hundred observations, in which 34.8% were censored. 37% were positive, while 28.2% were negative. The prior distributions are assumed as follows.

$$\begin{aligned} &\boldsymbol{\delta} \mid \tau^2 \sim \mathcal{N}\left(0, 10\tau^2 I_2\right), \quad \boldsymbol{\zeta} \mid \tau^2 \sim \mathcal{N}\left(0, 10\tau^2 I_2\right), \\ &\boldsymbol{\alpha} \mid \tau^2 \sim \mathcal{N}\left(0, 10\tau^2 I_2\right), \quad \tau^2 \sim \mathcal{IG}\left(0.1, 0.1\right). \end{aligned} \tag{13}$$



(c) Marginal posterior densities.

Fig. 1. (*continued*)

Table 3
Friction model: posterior means, standard deviations, 95% credible intervals, and inefficiency factors.

	True	Mean	Stdev	95% Interval	Inef
δ_1	−1.0	−0.990	0.105	(−1.195, −0.788)	106.8
δ_2	−0.4	−0.418	0.093	(−0.606, −0.247)	81.9
ζ_1	0.5	0.399	0.093	(0.222, 0.590)	48.6
ζ_2	0.2	0.233	0.077	(0.077, 0.381)	36.2
α_1	2.0	1.957	0.056	(1.851, 2.071)	34.9
α_2	0.5	0.499	0.038	(0.425, 0.574)	7.2
τ^2	0.49	0.519	0.047	(0.434, 0.618)	17.8

The proposed Gibbs sampler is implemented where we discard the initial 4000 samples and the subsequent 10,000 samples are recorded to conduct a Bayesian inference. Table 3 summarizes MCMC outputs, and reports that the posterior means are close to the true values, and all 95% credible intervals contain the true values. The inefficiency factors for $\boldsymbol{\zeta}, \boldsymbol{\alpha}$ and τ are not so large (7–49) while those for $\boldsymbol{\delta}$ are larger (80–110), suggesting that we may need a similar acceleration step to the Gibbs sampler as in Section 4.1.

Fig. 3 also shows the sample paths, sample autocorrelations functions, and estimated marginal posterior densities of the obtained MCMC samples. The sample paths indicate that the Markov chains are mixing well and all sample autocorrelations seem to decay fairly quickly to zero. The estimated marginal posterior densities have true values around their modes as expected.

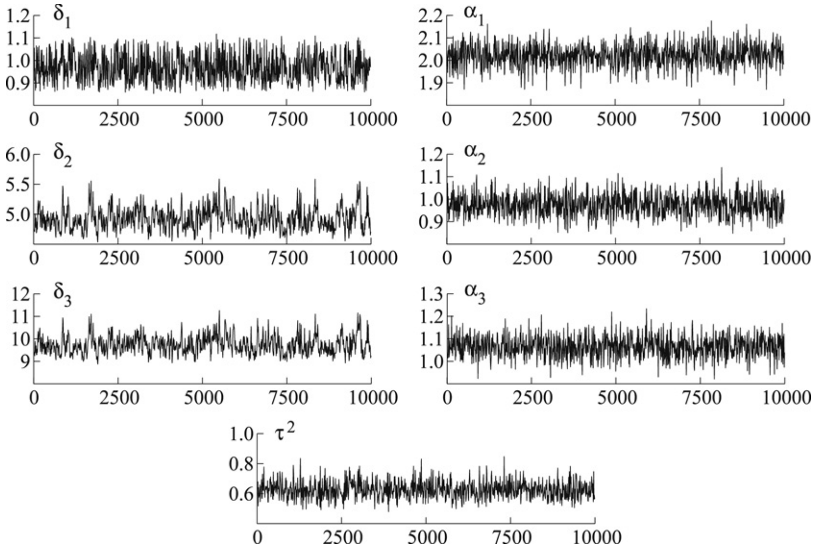
5. Empirical studies

5.1. Wage function of married women

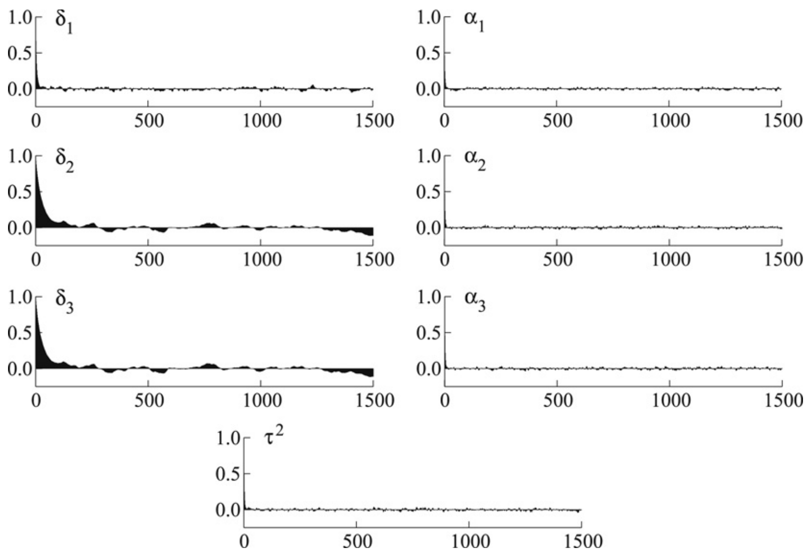
First we apply the Tobit model with covariate dependent thresholds to the wage function of married women using both of the simple and accelerated Gibbs samplers. We consider the popular labour supply data by Mroz (1987) on 753 married white women in 1975 and compare three candidate models:

Model $M1$: Type 1 Tobit model with a fixed threshold
Model $M2$: Type 1 Tobit model with covariate dependent thresholds
Model $M3$: Sample selection (Type 2 Tobit) model

using Deviance Information Criteria (DIC) rather than marginal likelihoods to remove the influences of the selected prior distributions. The ages of those women in the dataset are between thirty and sixty, and 428 individuals worked during the year (hence 43.2% of wage data are censored). Table 4 shows the variables considered in the following analysis.



(a) Sample paths.



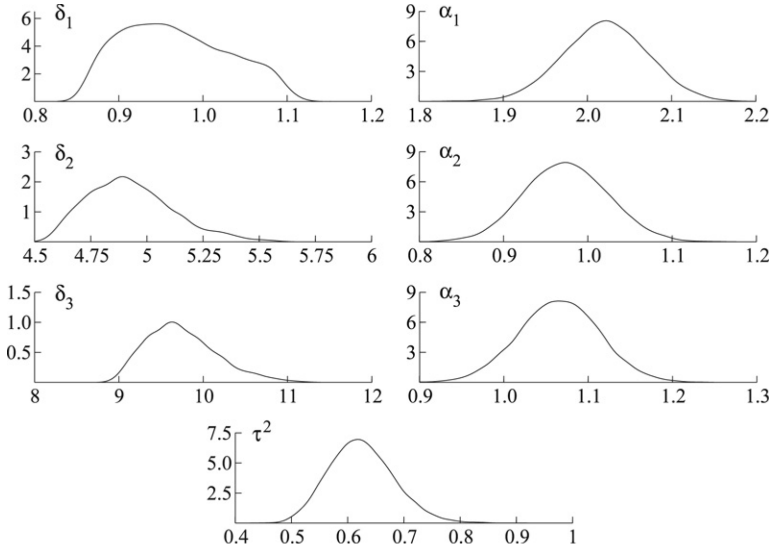
(b) Sample autocorrelations.

Fig. 2. Tobit model. Accelerated Gibbs sampler.

Table 4
Female labour supply data from Mroz (1987).

Variable name	Description
Wage	Wife's average hourly earnings (in 10 dollars)
Education	Years of schooling (standardized)
Age	Wife's age (standardized)
Income	Family income (standardized)
Kids	1 if there are children under 18, else 0
Experience	Actual years of wife's previous labour market experience (standardized)
City	1 if living in large urban area, else 0

The dependent variable is wife's average hourly earnings (in 10 dollars) and independent variables for the regression equation are education (years of schooling), experience (actual years of wife's previous labour), squared experience, and city (1 if living in large urban area, else 0). The education and experience variables are standardized such that their means and variances are equal to 0 and 1, respectively. The independent variables for the threshold equation and selection equation



(c) Marginal posterior densities.

Fig. 2. (*continued*)

Table 5
$M1$: Type 1 Tobit model with a fixed threshold. Posterior means, standard deviations, 95% credible intervals and inefficiency factors.

Variable	Mean	Stdev	95% Interval	Inef
Const	−6.546	1.045	(−8.601, −4.503)	1.2
Education	0.643	0.081	(0.484, 0.803)	1.0
Experience	2.167	0.240	(1.705, 2.647)	2.3
Experience^2	−0.593	0.138	(−0.866, −0.324)	1.8
City	−0.084	0.379	(−0.822, 0.661)	0.8
τ^2	20.013	1.483	(17.293, 23.108)	4.2

are education, age (wife's age), squared age, income (family income), and kids (1 if there are children under 18, else 0). The age and income variables are also standardized.

Model M1 (Type 1 Tobit model with a fixed threshold). We assume fairly flat prior distributions such that

$$\boldsymbol{\alpha}|\tau^2 \sim \mathcal{N}(0, 10\tau^2 I_5), \quad \tau^2 \sim \mathcal{IG}(0.1, 0.1).$$

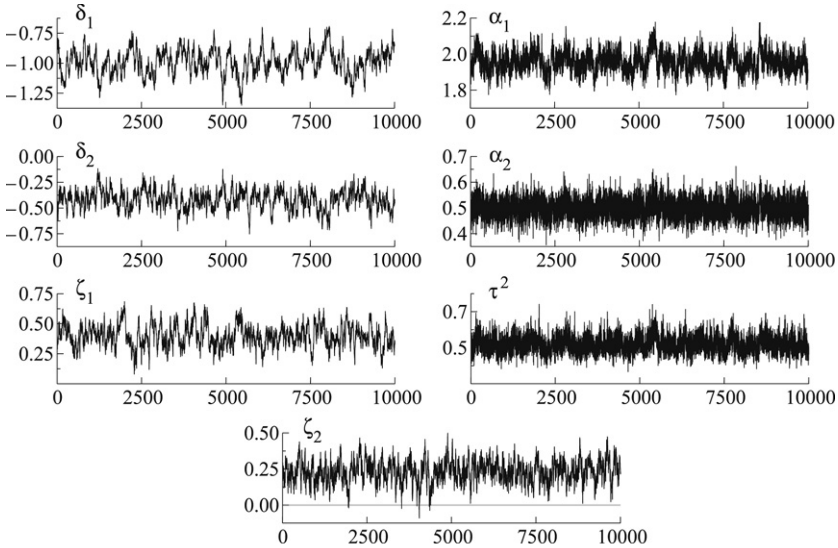
The initial 10,000 variates are discarded and the subsequent 50,000 values are recorded to conduct an inference. The summary statistics are given in Table 5. The inefficiency factors are very low, which implies the fast convergence of the distribution of MCMC samples to the posterior distribution. The 95% credible intervals do not include zero for variables, education, experience, and experience^2. The posterior probability of positive (negative) effects is greater than 0.95 for education and experience (for experience^2). This implies that the person with the higher education tends to receive the higher wage, while the effect of the experience on the wage is expected to be positive only for a limited length of the years of experience.

Model M2 (Type 1 Tobit model with covariate dependent thresholds). For the prior distributions, we use the same priors as in $M1$ and, further, assumed that $\boldsymbol{\delta}|\tau^2 \sim \mathcal{N}(0, 10\tau^2 I_6)$.

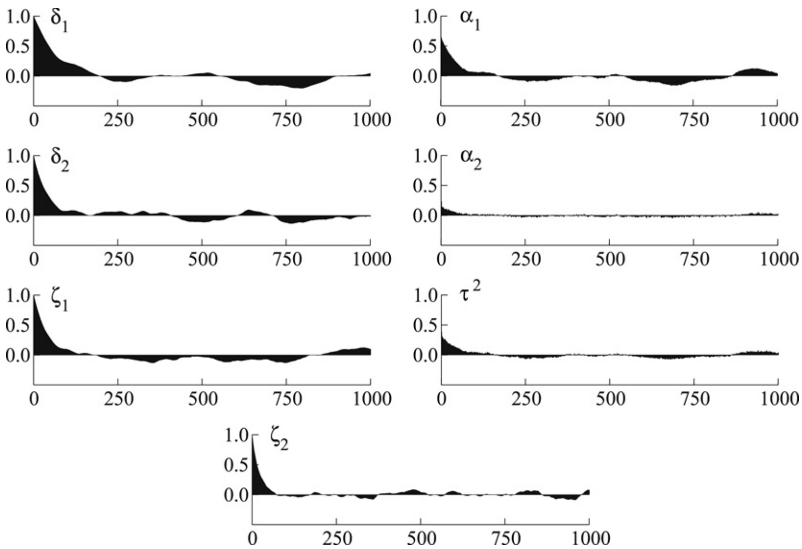
Using the generalised Gibbs sampler discussed in Section 2, the initial 20,000 variates are discarded and the subsequent 100,000 values are recorded where the acceptance rate of MH algorithm in the acceleration step was 57.6%. The summary statistics are given in Table 6. The inefficiency factors are based on the Gibbs sampler with the acceleration step (the factors without the acceleration step are given in brackets). The inefficiencies for regression equations are relatively small, while those for the threshold equation are still large in the range of 150–840 even with the acceleration step.

In the estimated threshold equation, the 95% credible intervals do not include zero for age^2 and income, indicating that the individual threshold of reserved wage varies with these characteristics. The married women with higher family income or older women tend to have higher reservation wage. For the regression equation, summary statistics are similar to those obtained in the standard Tobit model (Table 5).

Model M3 (Type 2 Tobit model). To implement the MCMC algorithms, we use reparameterizations, $\tau^2 = \phi + \gamma^2$ and $\rho = \gamma/\sqrt{\phi + \gamma^2}$ where τ^2 denotes the (2,2) element of Σ (see e.g., Omori (2007)). For the prior distributions, we assume that



(a) Sample paths.



(b) Sample autocorrelations.

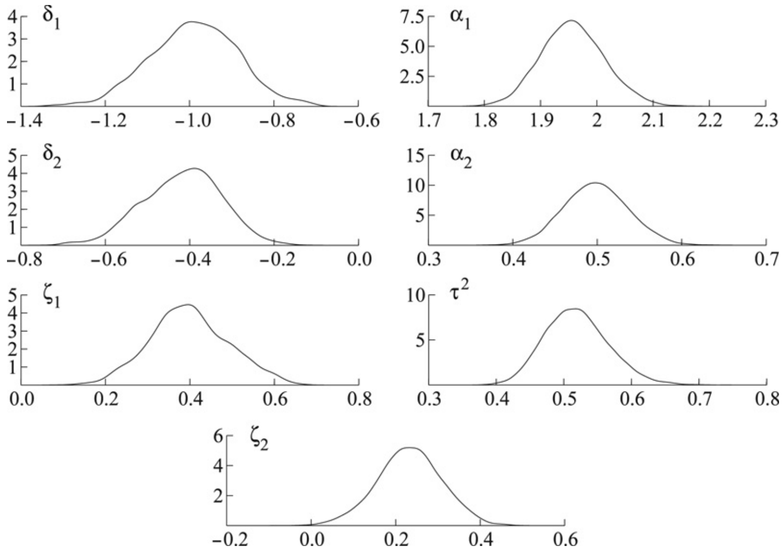
Fig. 3. Friction model.

$$\boldsymbol{\theta} \sim \mathcal{N}(0, 10I_6), \quad \boldsymbol{\beta} \sim \mathcal{N}(0, 10I_5), \quad \gamma \sim \mathcal{N}(0, 10), \quad \phi \sim \mathcal{IG}(0.1, 0.1).$$

The initial 10,000 variates are discarded and the subsequent 50,000 values are recorded. The summary statistics are given in Table 7. In the selection equation, the 95% credible intervals do not include zero for variables, education and income. The posterior probability that education (family income) has positive (negative) effects on the participation in the labour market is greater than 0.95. This indicates that the person with higher education or low family income tends to participate in the labour market.

For the regression equation, summary statistics are somewhat similar to those obtained in the standard Tobit model and a Tobit model with covariate dependent thresholds (Tables 5 and 6). The correlation coefficient ρ is found to be very high, suggesting that models $M1$ or $M2$ may be preferred.

We compared three models using DIC. The DIC and their standard errors are based on 20 iterations of 5000 samples, and are shown in Table 8. The model $M2$ attains the smallest value and selected as the best. It shows that our Tobit model with covariate dependent thresholds is a good alternative model to the standard Tobit and Type 2 Tobit models.



(c) Marginal posterior densities.

Fig. 3. (*continued*)

Table 6
M2: Type 1 Tobit model with covariate dependent thresholds Posterior means, standard deviations, 95% credible intervals and inefficiency factors [inefficiency factors obtained from the simple Gibbs sampler].

Variable	Mean	Stdev	95% Interval	Inef	[Inef]
Threshold equation					
Const	0.685	0.408	(−0.203, 1.385)	834.8	[1713.5]
Education	−0.039	0.028	(−0.090, 0.015)	819.1	[1748.8]
Age	0.114	0.056	(−0.003, 0.213)	166.5	[504.1]
Age^2	0.132	0.079	(0.015, 0.237)	153.2	[180.0]
Income	0.205	0.150	(0.072, 0.368)	403.0	[1084.5]
Kids	−0.027	1.008	(−0.295, 0.304)	434.1	[727.6]
Regression equation					
Const	−6.179	0.078	(−8.174, −4.221)	5.9	[7.1]
Education	0.623	0.230	(0.471, 0.778)	5.6	[7.0]
Experience	2.054	0.133	(1.610, 2.513)	3.0	[3.5]
Experience^2	−0.553	0.133	(−0.817, −0.295)	1.9	[2.1]
City	−0.030	0.365	(−0.749, 0.686)	0.6	[1.0]
τ^2	18.42	1.363	(15.927, 21.266)	6.8	[6.1]

Table 7
M3: Sample selection (Type 2 Tobit) model. Posterior means, standard deviations, 95% credible intervals and inefficiency factors.

Variable	Mean	Stdev	95% Interval	Inef
Selection equation				
Const	−1.618	0.246	(−2.100, −1.137)	14.8
Education	0.145	0.019	(0.107, 0.182)	13.9
Age	−0.014	0.028	(−0.071, 0.038)	61.7
Age^2	−0.000	0.026	(−0.052, 0.052)	37.2
Income	−0.067	0.029	(−0.121, −0.008)	53.4
Kids	−0.077	0.066	(−0.213, 0.048)	48.5
Regression equation				
Const	−6.320	0.931	(−8.164, −4.505)	3.8
Education	0.619	0.073	(0.476, 0.764)	2.4
Experience	0.398	0.123	(0.159, 0.642)	57.6
Experience^2	−0.078	0.072	(−0.225, 0.061)	25.1
City	0.040	0.194	(−0.343, 0.422)	41.0
σ^2	17.64	1.396	(15.035, 20.511)	83.5
ρ	0.991	0.004	(0.982, 0.996)	99.6

Table 8
Model comparison.

Model	$M1$	$M2$	$M3$
DIC (s.e.)	2937.4 (0.1)	2901.7 (0.8)	2984.8 (0.7)

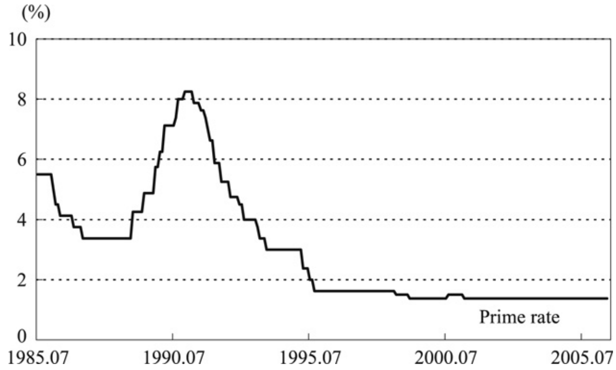


Fig. 4. Monthly prime rate data in Japan.

Table 9
Variables for the friction model of prime rate changes.

Variable name	Description
$dprime_t$	Difference of the prime rates at months t and $t-1$
$dcall_t$	Difference of the average call rates at months $t-1$ and $t-2$ (standardized)
$damount_t$	Difference of the outstanding amounts in the call money market at the end of months $t-1$ and $t-2$ (standardized)
iip_t	Index of the industrial productions for durable consumer goods shipment at month t (standardized)
$sales_t$	Commercial sales value ratio of the wholesale at month t: ratio to the same month of the previous year (standardized)
$unemploy_t$	Unemployment rate at month t (standardized)

5.2. Prime rate changes

The second example is an application of the friction model with covariate dependent thresholds to the prime rate changes in Japan. The prime rate is the bank's lending rate for most favorable customers, and, because of its sticky movement, a friction model is sometimes applied to analyze prime rate changes (see e.g., Forbes and Mayne (1989)). In this example, we use the monthly short-term prime rate used by principal banks from July 1985 through June 2006 reported by the Bank of Japan (see Fig. 4).

Table 9 summarizes the variables we considered. The dependent variable, y_t, is the difference of the prime rates at times t and $t-1$. There are two hundred and fifty observations in which 9 and 24 observations show positive and negative prime rate changes, respectively. The sample mean, standard deviation, minimum and maximum are −0.017, 0.173, −0.750, and 0.875.

The independent variables for the regression equation are differences of the average call rates and amounts (reported by the Bank of Japan) in the call money market at months $t-1$ and $t-2$, and independent variables for the threshold equations are selected from those related to the business cycle: index of industrial production for durable consumer goods shipment, commercial sales value ratio of the wholesale, and the unemployment rate at month t. First two variables are reported by Ministry of Economy, Trade and Industry, while the unemployment rate is reported by Ministry of Internal Affairs and Communications.

To estimate the parameters of the proposed friction model, we generated 100,000 MCMC samples after discarding 40,000 samples, The results are found in Table 10.

In the regression equation, the posterior probability that the call rate changes have a positive effect on prime rate changes is greater than 0.95. This result is quite natural because the call rate is used as one of the monetary policy instruments for the central bank in order to control interest rates.

Among independent variables for the threshold equations, the unemployment rate has a negative effect on the lower threshold, while other variables (iip and sales) have negative effects on the upper threshold, because their 95% credible intervals do not include 0. The lower unemployment rate would increase the lower threshold, while the higher index of industrial production or the higher commercial sales value ratio would decrease the upper threshold. Thus, these results indicate that banks would change the prime rates more often when the economy is expanding than when it is declining.

Table 10
Friction model of prime rate changes: Posterior means, standard deviations, 95% credible intervals and inefficiency factors.

	Mean	Stdev	95% Interval	Inef
Threshold equations				
Lower threshold				
δ_1 (const)	−1.746	0.441	(−2.846, −1.090)	359.9
δ_2 (iip)	−0.150	0.160	(−0.495, 0.140)	47.3
δ_3 (sales)	−0.187	0.180	(−0.571, 0.143)	45.0
δ_4 (unemploy)	−0.511	0.206	(−0.993, −0.183)	148.3
Upper threshold				
ζ_1 (const)	2.848	0.821	(1.659,4.845)	358.5
ζ_2 (iip)	−0.537	0.301	(−1.228, −0.052)	158.6
ζ_3 (sales)	−0.809	0.373	(−1.689, −0.235)	220.1
ζ_4 (unemploy)	−0.456	0.296	(−1.124, 0.054)	126.3
Regression equation				
α_1 (dcall)	0.327	0.118	(0.132, 0.598)	137.8
α_2 (damount)	0.180	0.129	(−0.049, 0.462)	57.6
τ^2	1.152	0.558	(0.484, 2.619)	348.2

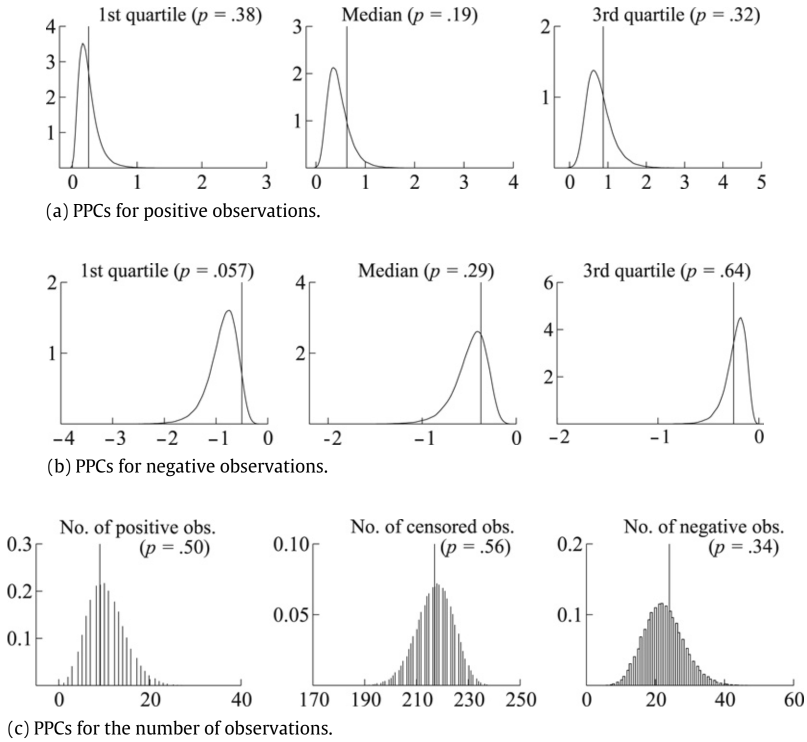


Fig. 5. Posterior predictive checks. Posterior predictive p-values are in parentheses.

Forbes and Mayne (1989) analyzed the rigidity of prime rate changes in the United States by the friction model, and their model is the special case of our proposed model, that is, thresholds include only unknown constants. They estimated model parameters by the maximum likelihood method for three consecutive periods. The estimated range between lower and upper thresholds, or the estimated range of no prime rate change, varies depending on the period. For the period October 1979 through October 1982, this range does not significantly differ from zero, while, for other periods (January 1977 through September 1979, and November 1982 through August 1987), it significantly differs from zero. Our further analysis for prime rate changes revealed that such a threshold variation would depend on the state of the economy.

At the end of this empirical study, we conducted the numerical posterior predictive checks (PPCs) to assess the plausibility of our friction model. PPC compares two test quantities, one based on observed data and the other based on replicated data from the predictive distribution. More detailed discussions are found, e.g., in Chapter 6 of Gelman et al. (2003).

We selected nine test quantities for the PPCs: three quartiles for positive and negative observations, and the number of positive, censored, and negative observations. PPC results are found in Fig. 5. The density plots and histograms are those of the test quantities based on replicated data and the vertical lines represent those based on observed data. We also calculated the posterior predictive p-values using the percentage of the replicated test quantities greater than the observed quantity, and showed them in parentheses.

These p-values indicate that all test quantities based on observed data are contained in the 95% intervals of those based on replicated data. Thus, posterior predictive checks suggest that our friction model is plausible to describe the prime rate changes' data. The first quartile for negative observations, however, shows small p-value, which would improve if we observe more negative prime rate changes.

6. Conclusion

A Bayesian analysis of a Tobit model with covariate dependent thresholds is described using the MCMC estimation method, and the acceleration step based on Liu and Sabatti (2000) is introduced to improve the mixing property of the MCMC samples. Two important extensions of the proposed models are also discussed.

Numerical examples using simulated data are given to illustrate the proposed estimation methods and their efficiencies are investigated. In empirical studies, using the labour supply data on the married women, the Tobit model with covariate dependent thresholds is estimated and model comparisons are conducted based on Deviance Information Criteria. Further, the friction model with covariate dependent thresholds is estimated using the prime rate data in Japan.

Acknowledgements

This work is supported by the Grants-in-Aid for Scientific Research 18330039 from the Japanese Ministry of Education, Science, Sports, Culture and Technology. The authors thank the Editor, the Associate editor, anonymous referees, Nobuhiro Yoneda, Professors Hideo Kozumi (Kobe University), Sylvia Frühwirth-Schnatter (Johannes Kepler Univertät Linz), and Charles S. Bos (Vrije Universiteit Amsterdam) for their helpful comments and discussions. All the computational results in Sections 4 and 5 are generated using Ox version 5.10 (see Doornik (2002)).

Appendix A

A.1. Posterior probability densities for a standard Tobit model

Joint posterior probability density of $(\boldsymbol{y}_c^*, \boldsymbol{\alpha}, \tau^2)$.

$$\pi(\boldsymbol{y}_c^*, \boldsymbol{\alpha}, \tau^2|\boldsymbol{y}_o) \propto (\tau^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{y}^* - X\boldsymbol{\alpha})'(\boldsymbol{y}^* - X\boldsymbol{\alpha})\right\} \prod_i I(y_{c,i}^* < d)$$
$$\times (\tau^2)^{-\frac{K}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{\alpha} - \boldsymbol{a}_0)'A_0^{-1}(\boldsymbol{\alpha} - \boldsymbol{a}_0)\right\} \times (\tau^2)^{-\left(\frac{n_0}{2}+1\right)} \exp\left\{-\frac{S_0}{2\tau^2}\right\}.$$

Conditional posterior probability density of $(\boldsymbol{\alpha}, \tau^2)$.

$$\pi(\boldsymbol{\alpha}, \tau^2|\boldsymbol{y}_c^*, \boldsymbol{y}_o) = \pi(\tau^2|\boldsymbol{y}_c^*, \boldsymbol{y}_o)\pi(\boldsymbol{\alpha}|\tau^2, \boldsymbol{y}_c^*, \boldsymbol{y}_o)$$
$$\propto (\tau^2)^{-\left(\frac{n_1}{2}+1\right)} \exp\left\{-\frac{S_1}{2\tau^2}\right\} \times (\tau^2)^{-\frac{K}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{\alpha} - \boldsymbol{a}_1)'A_1^{-1}(\boldsymbol{\alpha} - \boldsymbol{a}_1)\right\},$$

where $n_1 = n_0 + n$, $S_1 = \boldsymbol{y}^{*'}\boldsymbol{y}^* + S_0 + \boldsymbol{a}_0'A_0^{-1}\boldsymbol{a}_0 - \boldsymbol{a}_1'A_1^{-1}\boldsymbol{a}_1$, $A_1^{-1} = A_0^{-1} + X'X$, $\boldsymbol{a}_1 = A_1(A_0^{-1}\boldsymbol{a}_0 + X'\boldsymbol{y}^*)$.
Conditional posterior distribution of $\boldsymbol{y}_c^*$. $y_{c,i}^* \sim \mathcal{TN}_{(-\infty,d)}(\boldsymbol{x}_i'\boldsymbol{\alpha}, \tau^2)$ for censored observation.

A.2. Joint posterior density of $(\boldsymbol{y}_c^*, \boldsymbol{\alpha}, \boldsymbol{\delta}, \tau^2)$

$$\pi(\boldsymbol{y}_c^*, \boldsymbol{\alpha}, \boldsymbol{\delta}, \tau^2|\boldsymbol{y}_o) \propto (\tau^2)^{-\frac{n}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{y}^* - X\boldsymbol{\alpha})'(\boldsymbol{y}^* - X\boldsymbol{\alpha})\right\} \prod_i I(y_{c,i}^* < \boldsymbol{w}_i'\boldsymbol{\delta}) \prod_j I(y_{o,j}^* \geq \boldsymbol{w}_j'\boldsymbol{\delta})$$
$$\times (\tau^2)^{-\frac{K}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{\alpha} - \boldsymbol{a}_0)'A_0^{-1}(\boldsymbol{\alpha} - \boldsymbol{a}_0)\right\} \times (\tau^2)^{-\frac{J}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{\delta} - \boldsymbol{d}_0)'D_0^{-1}(\boldsymbol{\delta} - \boldsymbol{d}_0)\right\}$$
$$\times (\tau^2)^{-\left(\frac{n_0}{2}+1\right)} \exp\left\{-\frac{S_0}{2\tau^2}\right\}.$$

Conditional posterior probability density of $(\boldsymbol{\alpha}, \tau^2)$.

$$\pi(\boldsymbol{\alpha}, \tau^2|\boldsymbol{\delta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o) = \pi(\tau^2|\boldsymbol{\delta}, \boldsymbol{y}_c^*, \boldsymbol{y}_o) \times \pi(\boldsymbol{\alpha}|\boldsymbol{\delta}, \tau^2, \boldsymbol{y}_c^*, \boldsymbol{y}_o)$$
$$\propto (\tau^2)^{-\left(\frac{n_1}{2}+1\right)} \exp\left\{-\frac{S_1}{2\tau^2}\right\} \times (\tau^2)^{-\frac{K}{2}} \exp\left\{-\frac{1}{2\tau^2}(\boldsymbol{\alpha} - \boldsymbol{a}_1)'A_1^{-1}(\boldsymbol{\alpha} - \boldsymbol{a}_1)\right\},$$

where $n_1 = n_0 + n + J$, $S_1 = \boldsymbol{y}^{*'}\boldsymbol{y}^* + S_0 + \boldsymbol{a}_0'A_0^{-1}\boldsymbol{a}_0 - \boldsymbol{a}_1'A_1^{-1}\boldsymbol{a}_1 + (\boldsymbol{\delta} - \boldsymbol{d}_0)'D_0^{-1}(\boldsymbol{\delta} - \boldsymbol{d}_0)$, $A_1^{-1} = A_0^{-1} + X'X$, $\boldsymbol{a}_1 = A_1(A_0^{-1}\boldsymbol{a}_0 + X'\boldsymbol{y}^*)$.

Conditional posterior distribution of $\boldsymbol{y}_c^*$. $y_{c,i}^*|\boldsymbol{\alpha}, \boldsymbol{\delta}, \tau^2 \sim \mathcal{TN}_{(-\infty, \boldsymbol{w}_i'\boldsymbol{\delta})}(\boldsymbol{x}_i'\boldsymbol{\alpha}, \tau^2)$ for censored observation.
Conditional posterior distribution of $\boldsymbol{\delta}$. $\boldsymbol{\delta}|\tau^2, \boldsymbol{y}_c^*, \boldsymbol{y}_o \sim \mathcal{TN}_{R_o\cap R_c}(\boldsymbol{d}_0, \tau^2 D_0)$.

Appendix B. Acceleration of the Gibbs sampler

To accelerate the convergence, we use a generalised Gibbs sampler introduced by Liu and Sabatti (2000). Consider a scale group $\Gamma = \{g > 0 : g(\boldsymbol{\varphi}) = (g\boldsymbol{\delta}, g\boldsymbol{\alpha}, g\tau, g\boldsymbol{y}_c^*)\}$ where $\boldsymbol{\varphi} = (\boldsymbol{\delta}, \boldsymbol{\alpha}, \tau, \boldsymbol{y}_c^*)$. The unimodular left-Harr measure $L(dg)$ for this scale group is $L(dg) = g^{-1}dg$ and the corresponding Jacobian is $J_g = g^{J+K+1+n-m}$. Let $(\boldsymbol{w}_{o,i}, \boldsymbol{x}_{o,i})$ denote the independent variable vectors of the ith observed response $y_{o,i}$ and $X_o' = (\boldsymbol{x}_{o,1}, \ldots, \boldsymbol{x}_{o,m})$. By Theorem 1 of Liu and Sabatti (2000), the conditional probability density of g is given by

$$\begin{aligned}\pi(g|\boldsymbol{\varphi}, \boldsymbol{y}_o) &\propto \pi(g\boldsymbol{\delta}, g\boldsymbol{\alpha}, g\tau, g\boldsymbol{y}_c^*|\boldsymbol{y}_o) \times |J_g| \times L(\mathrm{d}g)\\ &\propto g^{-(n_0+m-1)} \times g^{-2}\exp\left[-\frac{-2qg^{-1}+rg^{-2}}{2\tau^2}\right] I(\lambda_L \le g^{-1} \le \lambda_U),\end{aligned}$$

where $q = \boldsymbol{\alpha}'X_o'\boldsymbol{y}_o + \boldsymbol{\alpha}'A_0^{-1}\boldsymbol{a}_0 + \boldsymbol{d}_0'D_0^{-1}\boldsymbol{\delta}$ and $r = \boldsymbol{y}_o'\boldsymbol{y}_o + \boldsymbol{a}_0'A_0^{-1}\boldsymbol{a}_0 + \boldsymbol{d}_0'D_0^{-1}\boldsymbol{d}_0 + S_0$. Since this is not a well-known distribution, we conduct Metropolis–Hastings algorithm to sample from $\pi(g|\boldsymbol{\varphi}, \boldsymbol{y}_o)$. When $\lambda_U < \infty$, we generate a candidate $g'^{-1} \sim \mathcal{U}(\lambda_L, \lambda_U)$ given the current point g. We accept g' with probability

$$\min\left[1, \left(\frac{g'}{g}\right)^{-(n_0+m-1)} \exp\left\{-\frac{-2q(g'^{-1}-g^{-1})+r(g'^{-2}-g^{-2})}{2\tau^2}\right\}\right].$$

We usually need to repeat the algorithm until it converges to $\pi(g|\boldsymbol{\varphi}, \boldsymbol{y}_o)$. However, by Theorem 2 of Liu and Sabatti (2000), we only need to conduct Metropolis–Hastings algorithm once using the initial value $g = 1$ since the Metropolis–Hastings transition kernel function satisfies $T_{\boldsymbol{\varphi}}(g, g') = T_{g_0^{-1}(\boldsymbol{\varphi})}(gg_0, g'g_0)$ for all $g, g', g_0 \in \Gamma$ where

$$T_{\boldsymbol{\varphi}}(g, g')L(dg') = \frac{I(\lambda_L < g'^{-1} < \lambda_U)}{g'^2(\lambda_U - \lambda_L)} \min\left[1, \left(\frac{g'}{g}\right)^{-(n_0+m-1)} \exp\left\{-\frac{-2q(g'^{-1}-g^{-1})+r(g'^{-2}-g^{-2})}{2\tau^2}\right\}\right] \mathrm{d}g',$$

where $g \neq g'$, and $T_{\boldsymbol{\varphi}}(g, g)L(dg) = 1 - \int T_{\boldsymbol{\varphi}}(g, g')L(dg')$. Thus, we generate a candidate g' and accept it with probability

$$\min\left[1, g'^{-(n_0+m-1)} \exp\left\{-\frac{-2q(g'^{-1}-1)+r(g'^{-2}-1)}{2\tau^2}\right\}\right].$$

Similarly, when $\lambda_U = \infty$, we generate a candidate $g'^{-1} \sim \mathcal{TN}_{(\lambda_L, \infty)}\left(1, \sigma_\lambda^2\right)$, for some σ_λ^2 and accept g' with probability

$$\min\left[1, g'^{-(n_0+m-1)} \exp\left\{-\frac{-2q(g'^{-1}-1)+r(g'^{-2}-1)}{2\tau^2} + \frac{(g'^{-1}-1)^2}{2\sigma_\lambda^2}\right\}\right].$$

References

Amemiya, T., 1984. Tobit models, a survey. Journal of Econometrics 24 (1–2), 3–61.
Charemza, W.W., Majerowska, E., 2000. Regulation of the Warsaw stock exchange: The portfolio allocation problem. Journal of Banking & Finance 24 (4), 555–576.
Chib, S., 1992. Bayes inference in the Tobit censored regression model. Journal of Econometrics 51 (1–2), 79–99.
Chib, S., 1995. Marginal likelihood from the Gibbs output. Journal of the American Statistical Association 90 (432), 1313–1321.
Chib, S., 2001. Markov chain Monte Carlo methods: Computation and inference. In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 5. North-Holland, Amsterdam, pp. 3569–3649.
Christofidesc, L.N., Li, D., 2005. Nominal and real wage rigidity in a friction model. Economics Letters 87 (2), 235–241.
Congdon, P., 2005. Bayesian predictive model comparison via parallel sampling. Computational Statistics & Data Analysis 48 (4), 735–753.
Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 39 (5), 829–844.
De Jong, M.G., Steenkamp, J.-B.E., Fox, J.-P., 2007. Relaxing measurement invariance in cross-national consumer research using a hierarchical IRT model. Journal of Consumer Research 34, 260–278.
Desarbo, W.S., Rao, V.R., Steckel, J.H., Wind, J., Colombo, R., 1987. A friction model for describing and forecasting price changes. Marketing Science 6 (4), 299–319.
Doornik, J.A., 2002. Object-Oriented Matrix Programming Using Ox, 3rd edition. Timberlake Consultants Press and Oxford, London.
Forbes, S.M., Mayne, L.S., 1989. A friction model of the prime. Journal of Banking and Finance 13 (1), 127–135.
Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, 2nd edition. Chapman & Hall/CRC, London.
Geweke, J., 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics, vol. 4, pp. 169–188, New York.
Hill, D.H., 1987. Derived demand estimation with survey experiments: Commercial electric vehicles. The Review of Economics and Statistics 69 (2), 277–285.
Johnson, T.R., 2003. On the use of heterogeneous thresholds ordinal regression models to account for individual differences in response style. Psychometrika 68 (4), 563–583.
King, G., Murray, C.J.L., Salomon, J.A., Tandon, A., 2004. Enhancing the validity and cross-cultural comparability of measurement in survey research. American Political Review 98 (1), 191–207.

Lin, T.-F., Schmidt, P., 1984. A test of the Tobit specification against an alternative suggested by Cragg. The Review of Economics and Statistics 66 (1), 174–177.

Liu, J.S., Sabatti, C., 2000. Generalised Gibbs sampler and multigrid Monte Carlo for Bayesian computation. Biometrika 87 (2), 353–369.

Maddala, G.S., 1983. Limited-Dependent and Qualitative Variables in Econometrics. In: Econometric Society Monographs, vol. 3. Cambridge University Press, Cambridge, New York.

Melenberg, B., Soest, A.V., 1996. Parametric and semi-parametric modelling of vacation expenditures. Journal of Applied Econometrics 11 (1), 59–76.

Mroz, T.A., 1987. The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions. Econometrica 55 (4), 765–799.

Omori, Y., 2007. Efficient Gibbs sampler for Bayesian analysis of a sample selection model. Statistics & Probability Letters 77 (12), 1300–1311.

Rosett, R.N., 1959. A statistical model of friction in economics. Econometrica 27 (2), 263–267.

Shriner, D., Yi, N., 2008. Deviance information criterion (DIC) in Bayesian multiple QTL mapping. Computational Statistics & Data Analysis 53 (5), 1850–1860.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., van der Linde, A., 2002. Bayesian measure of model complexity and fit. Journal of the Royal Statistical Society, Series B 64, 583–639.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica 26 (1), 24–36.

Zuehlke, T.W., 2003. Estimation of a Tobit model with unknown censoring threshold. Applied Economics 35 (10), 1163–1169.