行政院國家科學委員會專題研究計畫　成果報告

動態樣本選擇模型之貝氏分析

計畫類別：個別型計畫
計畫編號：NSC92-2415-H-032-009-
執行期間：92 年 08 月 01 日至 93 年 07 月 31 日
執行單位：淡江大學財務金融學系(所)

計畫主持人：黃河泉

報告類型：精簡報告

處理方式：本計畫可公開查詢

中　華　民　國 93 年 11 月 4 日

Bayesian inference of the dynamic sample selection model

Ho-Chuan (**River**) Huang [1]

October 30, 2004

Abstract

Sample selection models are often applied to cross-sectional data in analyzing decisions such as whether to participate in the labor market and if do, how to determine the desired number of labor hours supplied. In contrast, this paper focuses on the modeling strategy when time series data are used. The novel dynamic sample selection model is estimated via the Gibbs sampler with data augmentation algorithm. The practicality of our model and estimation technique is illustrated by using simulated data as well as real data on dividend payout to analyze the determinants of firms' whether-to-pay and how-much-to-pay decisions over time.

Keywords: dynamic sample selection model; Gibbs sampler; data augmentation; dividend

JEL classification: C11; C22; C25

1 Introduction

Limited dependent variable (LDV) models are regression models in which the dependent variable is limited in some ways, e.g., truncated or censored. Among which, one of the most popular models is known as the Tobit model, first introduced in a pioneering work by Tobit (1958). In the work, Tobin (1958) investigates the effect of household expenditure on durable goods by explicitly taking into account the fact that the expenditure cannot be negative. In a similar spirit, Kim and Maddala (1992) employ a Tobit model to analyze the dividend payout decisions of firms by recognizing the fact that some firms may not always pay dividends ('zero' dividend) and same firm may not always pay dividend over time.

Another widely-used LDV model is the sample selection (also known as selection bias, or Type 2 Tobit) model. One famous example is the female labor supply which typically consists of two equations. The first (participation) equation decides whether or not to work while the second (wage) equation, given that she chooses to participate in the labor market, determines what characteristics might affect the wage rates. Please see Heckman (1974,

[1]Corresponding author: **Ho-Chuan (River) Huang**, Department of Banking and Finance, Tamkang University, 151 Ying-Chun Road, Tamsui 25137, Taipei County, **Taiwan**. Email: **river@mail.tku.edu.tw**. Website: **http://mail.tku.edu.tw/river/**. Financial support from NSC, Taiwan under grant # NSC 92-2415-H-032-009 is greatly appreciated. Any remaining errors are my own responsibility.

1979) for the theoretical motivation and empirical estimation of such model. Recently, in a paper related to Kim and Maddala (1992), Huang (2001) implements the sample selection model to analyze the firms' dividend decisions. In particular, Huang (2001) allows for the possibility that companies first decide whether to pay dividend by a selection regression and, conditional on the companies have decided to pay, how much dividend will be paid via the outcome regression. As in Scollnik (1993), Huang (2001) argues that such modeling strategy allows for the eventuality that these two decisions might be affected by different variables and a given variable might influence each of the two decisions differentially.

Note that, in most studies, the above Tobit or/and sample selection models are analyzed using either cross-section data or panel data, e.g., Maddala (1983) and Amemiya (1985). Recently, those LDV-related models have been increasingly used in analyzing time series data. For instance, Dueker (1999) adds Markov-switching heteroskedasticity to a dynamic ordered probit regression to address the issue of discrete change in the bank prime lending rate while Dueker (2005) presents a new Qual VAR model for incorporating information from qualitative variables (discrete data) in vector autoregressive models to predict the 2001 recession out of sample. Alternatively, Zangari (1994) analyzes the Federal Reserve's holdings of securities under repurchase agreements by using the Tobit model with autoregressive errors while Zangari and Tsurumi (1996) utilize similar framework to analyze the exports of Japanese passenger cars to the United States under a voluntary export restraint. Similarly, Wei (1999) proposes a Bayesian approach via the Gibbs sampler to estimate dynamic Tobit models and Wei (2002) considers a censored-GARCH model to analyze the Treasury bill futures returns process with price limits. Recently, Huang (2003) tests the capital asset pricing model by allowing for regime-switching risks and taking account of price limit regulation.

This paper, in a similar spirit, generalizes the sample selection model to analyze time series data. The novel model is called "dynamic sample selection" (DSS) model which might appear more appropriate in many cases. For example, in the finance literature, it is found as an empirical regularity that dividend ratios are significant predictors of annual equity premia, e.g., Cochrane (1997). As a result, knowing the decisions of dividend payout for each company (over time) could provide useful information for how investors should divide their assets between stocks and bonds, based on the premise that equity premia vary in a predictable (by the dividend and other variables) fashion, e.g., Campbell and Viceira

(1999) and Campbell, Chan and Viceira (2003). Therefore, a modification of Huang's (2001) sample selection model has to be made in order to take into account firm's dividend payout decisions over time. This DSS model, to the best of our knowledge, has not been discussed or analyzed in the literature, either in econometrics or finance.

However, the likelihood function for the DSS model is complicated and difficult to evaluate due to the need of multiple integration. Alternatively, a practical simulation-based approach via the Markov chain Monte Carlo, e.g., Gibbs sampler or/and Metropolis-Hastings algorithm with data augmentation, is developed to overcome the problems in fitting the DSS model. Our Bayesian approach is shown to be both conceptually simple and computationally feasible. In addition, the estimates exhibit exact finite-sample properties which are also desirable in our application with rather limited observations.

2 The dynamic sample selection model

Consider the dynamic ('p'th-order autoregressive) sample selection model,

$$y_{1t}^* = \mathbf{y}_{1t}^{*'}\phi_1 + x_{1t}'\beta_1 + \epsilon_{1t} \quad (1)$$

$$y_{2t}^* = \mathbf{y}_{2t}^{*'}\phi_2 + x_{2t}'\beta_2 + \epsilon_{2t} \quad (2)$$

where, for $j = 1, 2$,

$$\begin{aligned} \mathbf{y}_{jt}^* &= (y_{j,t-1}^*, y_{j,t-2}^*, \cdots, y_{j,t-p}^*)' \\ \phi_j &= (\phi_{j1}, \phi_{j2}, \cdots, \phi_{jp})' \\ x_{jt} &= (x_{jt,1}, x_{jt,2}, \cdots, x_{jt,k})' \\ \beta_j &= (\beta_{j1}, \beta_{j2}, \cdots, \beta_{jk})' \end{aligned}$$

The selection regression in (1) decides whether or not a company would pay dividend over time; if it does, the outcome regression in (2) will determine the amount of dividend the company will pay. The observed variable y_{1t} takes on the value 1 or 0 according to the positivity of the latent variable y_{1t}^*, i.e.,

$$y_{1t} = \begin{cases} 1 & \text{if } y_{1t}^* > 0 \\ 0 & \text{if } y_{1t}^* \leq 0 \end{cases} \quad (3)$$

In other words, we observe that the company decides to pay dividends if y_{1t}^* is positive (i.e., the intention to pay dividend is relatively large); otherwise, the company would not pay any dividend (i.e., the intention to pay dividend is relatively small). The desired value of dividend, y_{2t}^*, is observed if and only if a company is observed to pay dividend, i.e.,

$$y_{2t} = \begin{cases} y_{2t}^* & \text{if } y_{1t} = 1 \\ \text{unobservable} & \text{if } y_{1t} = 0 \end{cases} \tag{4}$$

otherwise, the desired value remains unobservable and recorded as 0 since institutional factors restrict the dividend to be non-negative. Note that only the sign of y_{1t}^* is observed and y_{2t}^* is observed only when $y_{1t}^* > 0$. Also, the explanatory variables x_{1t} and x_{2t} can be different. In particular, the inclusion of lagged latent dependent variables as additional explanatory variables makes the model dynamic.

The error term $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})'$ is assumed to be normally distributed as,

$$\begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix} \right] \tag{5}$$

where the variance of ϵ_{1t} is normalized to be 1 for identification reason.

We can rewrite equation (2) as,

$$\epsilon_{2t} = \sigma_{12}\epsilon_{1t} + \eta_t \tag{6}$$

where $\eta_t \sim iid\mathcal{N}(0, \sigma^2)$ and $\sigma^2 = \sigma_{22} - \sigma_{12}^2$ and re-parameterize the variance-covariance matrix as,

$$\Sigma = \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma^2 + \sigma_{12}^2 \end{pmatrix} \tag{7}$$

Our focus is now on the simulation of (σ^2, σ_{12}) since σ_{22} can be obtained by $\sigma^2 + \sigma_{12}^2$. It will become clear that such reformulations greatly simplify the analysis.

3 Posterior inference

Given the latent data $y_j^* = \left(y_{j1}^*, y_{j2}^*, \cdots, y_{jT}^*\right), j = 1, 2$, equations (1) and (2) can be written as,

$$y_t^* = X_t\gamma + \epsilon_t \tag{8}$$

where $y_t^* = (y_{1t}^*, y_{2t}^*)'$, $\gamma = (\phi_1', \beta_1', \phi_2', \beta_2')'$, $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})'$, and

$$X_t = \begin{pmatrix} \mathbf{y}_{1t}^{*\prime} & x_{1t}' & 0 & 0 \\ 0 & 0 & \mathbf{y}_{2t}^{*\prime} & x_{2t}' \end{pmatrix}$$

From (8), the (complete-data) likelihood function is,

$$(2\pi)^{-\frac{2(T-p)}{2}} \left|\Sigma^{-1}\right|^{\frac{(T-p)}{2}} \exp\left\{ -\frac{1}{2} \sum_{t=p+1}^{T} (y_t^* - X_t\gamma)'\Sigma^{-1}(y_t^* - X_t\gamma) \right\} \tag{9}$$

In order to complete the framework for Bayesian inference, we assume that the priors are,

$$\pi(\gamma, \sigma_{12}, \sigma^{-2}) \quad = \quad \pi(\gamma)\pi(\sigma_{12})\pi(\sigma^{-2})$$

i.e., all priors are independent. To be more specific, the prior of γ is normally distributed as $\pi(\gamma) \sim \mathcal{N}(\gamma_0, G_0)$, the prior of the covariance term σ_{12} is also normal as $\pi(\sigma_{12}) \sim \mathcal{N}(r_0, s_0)$ and the prior of σ^{-2} is gamma distributed as $\pi(\sigma^{-2}) \sim \mathcal{G}(\nu_0, \delta_0)$.

3.1 Full conditional distributions

It is clear that the main difficulty in estimating the model is due to the presence of the latent values $y^* = \left(y_1^{*'}, y_2^{*'}, \cdots, y_T^{*'}\right)'$. Therefore, we augment the data space by drawing the latent values from their respective full conditional density and, given the latent data, derive the full conditionals of the model parameters γ, σ_{12} and σ^{-2}.

1. Combined with the prior $\gamma \sim \mathcal{N}(\gamma_0, G_0)$, the full conditional density of γ is:

$$\gamma | y^*, \theta_{\setminus\gamma} \sim \mathcal{N}(\gamma_n, G_n) \tag{10}$$

where

$$\gamma_n \quad = \quad G_n \left(G_0^{-1}\gamma_0 + \sum_{t=p+1}^{T} X_t' \Sigma^{-1} y_t^* \right) \tag{11}$$

$$G_n \quad = \quad \left[G_0^{-1} + \sum_{t=p+1}^{T} X_t' \Sigma^{-1} X_t \right]^{-1} \tag{12}$$

2. Given y^* and γ, ϵ_{1t} and ϵ_{2t} are treated as known since they can be obtained by

$$\epsilon_{1t} \quad = \quad z_{1t} - \mathbf{z}_{1t}'\phi_1 - x_{1t}'\beta_1$$

$$\epsilon_{2t} \quad = \quad z_{2t} - \mathbf{z}_{2t}'\phi_2 - x_{2t}'\beta_2$$

Combining the prior of $\sigma_{12} \sim \mathcal{N}(r_0, s_0)$ with (6), i.e.,

$$(2\pi)^{-\frac{T-p}{2}} (\sigma^{-2})^{\frac{T-p}{2}} \exp\left\{ -\frac{1}{2\sigma^2} (\epsilon_2 - \sigma_{12}\epsilon_1)'(\epsilon_2 - \sigma_{12}\epsilon_1) \right\} \tag{13}$$

we can derive the full conditional density of σ_{12} to be,

$$\sigma_{12} | z, \theta_{\setminus\sigma_{12}} \sim \mathcal{N}(r_n, s_n) \tag{14}$$

where

$$r_n = s_n \left(s_0^{-1} r_0 + \epsilon_1' \epsilon_2 / \sigma^2\right) \tag{15}$$

$$s_n = \left(s_0^{-1} + \epsilon_1' \epsilon_1 / \sigma^2\right)^{-1} \tag{16}$$

3. Similarly, combining the prior of $\sigma^{-2} \sim \mathcal{G}(\nu_0, \delta_0)$ with the likelihood from (13), the full conditional density of σ^{-2} is,

$$\sigma^{-2} | y^*, \theta_{\backslash \sigma^{-2}} \sim \mathcal{G}(\nu_n, \delta_n) \tag{17}$$

where

$$\nu_n = \nu_0 + \frac{n}{2} \tag{18}$$

$$\delta_n = \delta_0 + \frac{(\epsilon_2 - \sigma_{12}\epsilon_1)'(\epsilon_2 - \sigma_{12}\epsilon_1)}{2} \tag{19}$$

4. The full conditional of y_1^*:

   It is clear that the key point to this model is the simulation of the latent data $y^* = (y_1^{*'}, y_2^{*'})'$. As can be shown, the full conditional density of y_{1t}^* is,

$$y_{1t}^* | y_{it} = 1, y_{1,\backslash t}^*, y_{2t}^*, \theta \sim \mathcal{TN}_{(0,\infty)} \left(\bar{\mu}_{1t}, \bar{\sigma}_1^2\right) \tag{20}$$

$$y_{1t}^* | y_{it} = 0, y_{1,\backslash t}^*, y_{2t}^*, \theta \sim \mathcal{TN}_{(-\infty,0])} \left(\bar{\mu}_{1t}, \bar{\sigma}_1^2\right) \tag{21}$$

   The y_{1t}^* is sampled from the range $(0, \infty)$ when $y_{1t} = 1$ and from $(-\infty, 0]$ when $y_{1t} = 0$.

5. The full conditional of y_2^*:

   In a similar way, it can be shown that the full conditional density of y_{2t}^* is,

$$y_{2t}^* | y_{1t}^*, y_{2,\backslash t}^*, \theta \sim \mathcal{N} \left(\bar{\mu}_{2t}, \bar{\sigma}_2^2\right) \tag{22}$$

4 Empirical results

For illustration, we consider a first-order dynamic sample selection model, i.e., $p = 1$, to analyze the firm's decisions on dividend payouts over time. The data of dividends, including both cash and stock dividends, and earnings per share (*eps*) of the Grape King Inc. (code number 1707) are taken from the AREMOS databank, Ministry of Education, Taiwan. In particular, the sample period starts from 1982 and ends in 2003, with a total of 22 annual observations. The relevant information is provided in Figure 1.

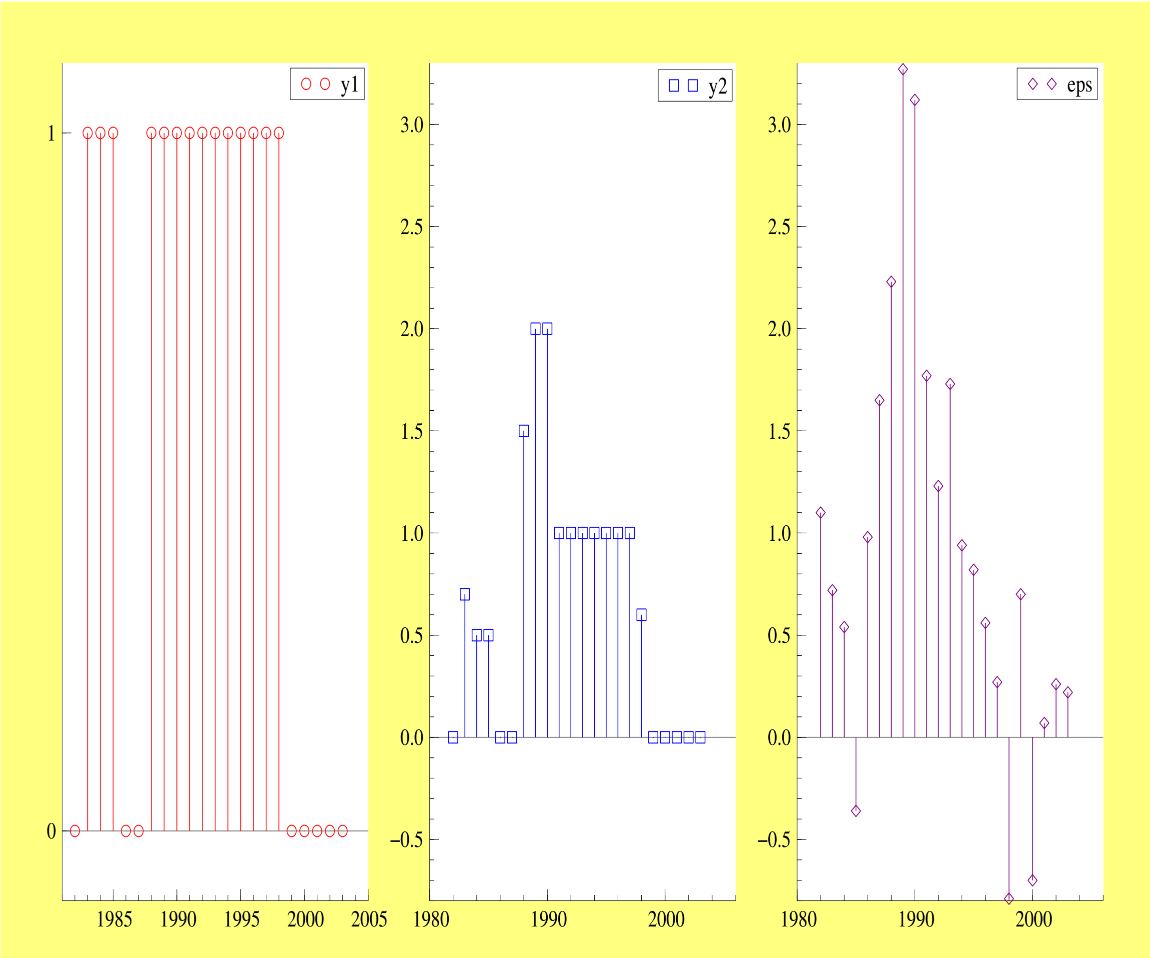

Figure 1: The data of y_1 (left panel; 1 denotes that the company pays dividends and 0 means that the company doesn't pay any dividend), y_2 (middle panel; the amount of dividend paid by the company) and *eps* (right panel; the earnings per share), respectively.

By applying the Gibbs sampler for 10,000 iterations, after discarding the first 10,000 random variates to mitigate the effect of initial values, the relevant moments and percentiles of the parameters are summarized in Table 1.

Table 1: The results of the first-order dynamic sample selection model

	mean	std	median	2.5%	5%	95%	97.5%
ϕ_1	0.7523	0.2054	0.8059	0.1764	0.3249	0.9802	1.0261
β_{10}	−0.5461	0.4738	−0.5347	−1.5109	−1.3539	0.1995	0.3366
β_{11}	0.8552	0.4474	0.8065	0.0993	0.1925	1.6775	1.8075
ϕ_2	0.1381	0.2176	0.1399	−0.2991	−0.2234	0.4899	0.5687
β_{20}	0.5813	0.2037	0.5749	0.1943	0.2548	0.9331	1.0067
β_{21}	0.3172	0.0893	0.3198	0.1310	0.1689	0.4596	0.4889
σ_{12}	−0.1385	0.1402	−0.1699	−0.3578	−0.3236	0.1474	0.2094
σ_{22}	0.0690	0.0387	0.0592	0.0268	0.0301	0.1395	0.1651

† The results are based on 10,000 random variates from Gibbs sampler after discarding the first 10,000 iterations.

It can be seen that the AR(1) coefficient in the (first) selection equation, i.e., ϕ_1, has a posterior mean of 0.7523 with standard deviation 0.2054. According to either 90% or 95% Bayesian confidence intervals, it is significantly positive. In other words, this evidence indicates that the intention of firm's whether-to-pay decision is persistent over time. In contrast, the AR(1) parameter in the (second) outcome equation, i.e., ϕ_2, has a positive posterior mean of 0.1381 with standard deviation 0.2176. Clearly, it is not significantly different from zero judged by either 90% or 95% Bayesian confidence intervals. Unlike the result in the selection equation, we find no evidence in support that the firm's (optimal) desirable amount of dividend payouts are related between adjacent time periods. However, the effects of earnings per share (*eps*) on the decisions of whether-to-pay and how-much-to-pay, measured by the coefficients $\beta_{11} = 0.8552$ and $\beta_{21} = 0.3172$ (posterior means), are both significantly positive relative to the Bayesian 90% or 95% confidence intervals.

5 Conclusions

This paper proposes a novel dynamic sample selection model and offers a Bayesian simulation-based estimation procedure for making posterior inference. In particular, we derive all the relevant full conditional densities required for implementing the Gibbs sampler. For illustration of its practical use, we apply the model to analyze the firm's whether-to-pay and how-much-to-pay decisions over time. Based on the annual sample from 1982 to 2003 of Grape King Inc., empirical results show that the decisions of whether-to-pay are correlated over time but the decisions of how-much-to-pay are not. In addition, as expected, the earnings per share variable affects positively the desires of whether-to-pay and how-much-to-pay decisions.

Appendix

Note that,

$$\begin{pmatrix} y^*_{1t} \\ y^*_{2t} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mathbf{y}^{*'}_{1t}\phi_1 + x'_{1t}\beta_1 \\ \mathbf{y}^{*'}_{2t}\phi_2 + x'_{2t}\beta_2 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{pmatrix}\right] \tag{23}$$

Then,

$$y^*_{1t}|y^*_{2t} \sim \mathcal{N}\left(\mu_{1t}, \omega_{11}\right) \tag{24}$$

where,

$$\begin{aligned} \mu_{1t} &= \mathbf{y}^{*'}_{1t}\phi_1 + x'_{1t}\beta_1 + \frac{\sigma_{12}}{\sigma_{22}}\left(y^*_{2t} - \mathbf{y}^{*'}_{2t}\phi_2 - x'_{2t}\beta_2\right) \\ \omega_{11} &= 1 - \frac{\sigma^2_{12}}{\sigma_{22}} \end{aligned}$$

Or, we can re-write (24) as,

$$y^*_{1t} = \phi_{11}y^*_{1,t-1} + \phi_{12}y^*_{1,t-2} + \cdots + \phi_{1p}y^*_{1,t-p} + x'_{1t}\beta_1 + \frac{\sigma_{12}}{\sigma_{22}}\epsilon_{2t} + \eta_{1t} \tag{25}$$

where $\eta_{1t} \sim \mathcal{N}(0, \sigma^2_{\eta_1})$ with $\sigma^2_{\eta_1} = 1 - \sigma^2_{12}/\sigma_{22}$.

Let $N = \min(T, t+p)$. Conditioning upon the parameters and all the other complete data, the relevant full conditional distribution of y^*_{1t} is:

$$\begin{aligned} & p(y^*_{1t}|y^*_{1\backslash t}, y^*_2, \theta) \\ = \; & p(y^*_{1t}|y^*_{11}, \cdots, y^*_{1t-1}, y^*_{1t+1}, \cdots, y^*_{1T}) \\ \propto \; & p(y^*_{1t}, y^*_{1t+1}, \cdots, y^*_{1T}|y^*_{11}, y^*_{12}, \cdots, y^*_{1t-1}) \\ \propto \; & p(y^*_{1t}, y^*_{1t+1}, \cdots, y^*_{1N}|y^*_{1t-1}, y^*_{1t-2}, \cdots, y^*_{1t-p}) \\ \propto \; & p(y^*_{1t}|y^*_{1t-1}, y^*_{1t-2}, \cdots, y^*_{1t-p}) \\ & \times p(y^*_{1t+1}|y^*_{1t}, y^*_{1t-1}, \cdots, y^*_{1t-p+1}) \\ & \times \cdots \times p(y^*_{1N}|y^*_{1N-1}, y^*_{1N-2}, \cdots, y^*_{1N-p}) \\ \propto \; & \exp\left\{-\frac{1}{2\sigma^2_{\eta_1}}\left(y^*_{1t} - \phi_{11}y^*_{1t-1} - \cdots - \phi_{1p}y^*_{1t-p} - x'_{1t}\beta_1 - \frac{\sigma_{12}}{\sigma_{22}}\epsilon_{2t}\right)^2\right\} \\ & \times \exp\left\{-\frac{1}{2\sigma^2_{\eta_1}}\left(y^*_{1t+1} - \phi_{11}y^*_{1t} - \cdots - \phi_{1p}y^*_{1t-p+1} - x'_{1t+1}\beta_1 - \frac{\sigma_{12}}{\sigma_{22}}\epsilon_{2t+1}\right)^2\right\} \\ & \times \cdots \times \exp\left\{-\frac{1}{2\sigma^2_{\eta_1}}\left(y^*_{1N} - \phi_{11}y^*_{1N-1} - \cdots - \phi_{1N-t}y^*_{1t} - x'_{1N}\beta_1 - \frac{\sigma_{12}}{\sigma_{22}}\epsilon_{2N}\right)^2\right\} \\ \propto \; & \exp\left\{-\frac{1}{2\sigma^2_{\eta_1}}\left[y^{*2}_{1t} - 2(w_{t,1})y^*_{1t}\right]\right\} \end{aligned}$$

$$\times \exp\left\{-\frac{1}{2\sigma_{\eta_1}^2}\left[\phi_{11}^2 y_{1t}^{*2} - 2(\phi_{11}w_{t,2})y_{1t}^*\right]\right\}$$
$$\times \cdots \times \exp\left\{-\frac{1}{2\sigma_{\eta_1}^2}\left[\phi_{1N-t}^2 y_{1t}^{*2} - 2(\phi_{1N-t}w_{t,N-t+1})y_{1t}^*\right]\right\}$$

where

$$\begin{aligned} w_{1t,1} &= \phi_{11}y_{1t-1}^* + \cdots + \phi_{1p}y_{1t-p}^* + x_{1t}'\beta_1 + \frac{\sigma_{12}}{\sigma^2+\sigma_{12}^2}\epsilon_{2t} \\ w_{1t,i} &= y_{1t+i-1}^* - \sum_{j=1,j\neq i-1}^{p}\phi_{1j}y_{1t+i-1-j}^* - x_{1t+i}'\beta_1 - \frac{\sigma_{12}}{\sigma^2+\sigma_{12}^2}\epsilon_{2t+i} \end{aligned}$$

and $t = 2, 3, \cdots, N-t+1$.

By completing squares, we can simplify the above equation to obtain,

$$\begin{aligned} &\exp\left\{-\frac{1}{2\left(\frac{\sigma_{\eta_1}^2}{1+\phi_{11}^2+\cdots+\phi_{1N-t}^2}\right)}\left[y_{1t}^* - \frac{w_{1t,1}+\phi_{11}w_{1t,2}+\cdots+\phi_{1N-t}w_{1t,N-t+1}}{1+\phi_{11}^2+\cdots+\phi_{1N-t}^2}\right]^2\right\} \\ = \; &\exp\left\{-\frac{1}{2\bar{\sigma}_1^2}(y_{1t}^*-\bar{\mu}_{1t})^2\right\} \end{aligned}$$

where the mean and variance are,

$$\bar{\mu}_{1t} = \frac{w_{1t,1}+\phi_{11}w_{1t,2}+\cdots+\phi_{1N-t}w_{1t,N-t+1}}{1+\phi_{11}^2+\cdots+\phi_{1N-t}^2} \tag{26}$$

$$\bar{\sigma}_1^2 = \frac{\sigma_{\eta_1}^2}{1+\phi_{11}^2+\cdots+\phi_{1N-t}^2} \tag{27}$$

We know that,

$$y_{2t}^*|y_{1t}^* \sim \mathcal{N}(\mu_{2t}, \omega_{22}) \tag{28}$$

where,

$$\begin{aligned} \mu_{2t} &= \mathbf{y}_{2t}^{*'}\phi_2 + x_{2t}'\beta_2 + \sigma_{12}\left(y_{1t}^* - \mathbf{y}_{1t}^{*'}\phi_1 - x_{1t}'\beta_1\right) \\ \omega_{22} &= \sigma^2 \end{aligned}$$

and we can re-write (28) as,

$$y_{2t}^* = \phi_{21}y_{2,t-1}^* + \phi_{22}y_{2,t-2}^* + \cdots + \phi_{2p}y_{2,t-p}^* + x_{2t}'\beta_2 + \sigma_{12}\epsilon_{1t} + \eta_{2t} \tag{29}$$

where $\eta_{2t} \sim \mathcal{N}(0, \sigma_{\eta_2}^2)$ with $\sigma_{\eta_2}^2 = \sigma^2$.

As above, the relevant full conditional distribution of y_{2t}^* is:

$$y_{2t}^* \sim \mathcal{N}(\bar{\mu}_{2t}, \bar{\sigma}_2^2) \tag{30}$$

where

$$\bar{\mu}_{2t} = \frac{w_{2t,1} + \phi_{21} w_{2t,2} + \cdots + \phi_{2N-t} w_{2t,N-t+1}}{1 + \phi_{21}^2 + \cdots + \phi_{2N-t}^2} \tag{31}$$

$$\bar{\sigma}_2^2 = \frac{\sigma_{\eta_2}^2}{1 + \phi_{21}^2 + \cdots + \phi_{2N-t}^2} \tag{32}$$

and, for $t = 2, 3, \cdots, N - t + 1$,

$$\begin{aligned} w_{2t,1} &= \phi_{21} y_{2t-1}^* + \cdots + \phi_{2p} y_{2t-p}^* + x_{2t}' \beta_2 + \sigma_{12} \epsilon_{1t} \\ w_{2t,i} &= y_{2t+i-1}^* - \sum_{j=1, j \neq i-1}^{p} \phi_{2j} y_{2t+i-1-j}^* - x_{2t+i}' \beta_2 - \sigma_{12} \epsilon_{1t+i} \end{aligned}$$

Reference

Albert, J. H. and Chib, S. (1993), "Bayesian analysis of binary and polychotomous response data." *Journal of the American Statistical Association* 88, 669-679.

Amemiya, T. (1985), *Advances Econometrics.* Harvard University Press, Cambridge.

Campbell, J. Y., Chan, Y. L. and Viceira, L. M. (2003), "A multivariate model of strategic asset allocation." *Journal of Financial Economics* 67, 41-80.

Campbell, J. Y. and Viceira, L. M. (1999), "Consumption and portfolio decisions when expected returns are time varying." *Quarterly Journal of Economics* 114, 433-495.

Casella, G. and George, E. I. (1992), "Explaining the Gibbs sampler." *American Statistician* 46, 167-174.

Chib, S. (1992), "Bayes inference in the Tobit censored regression model." *Journal of Econometrics* 51, 79-99.

Chib, S. (1995), "Marginal likelihood from the Gibbs output." *Journal of the American Statistical Association* 90, 1313-1321.

Chib, S. and Greenberg, E. (1995), "Hierarchical analysis of SUR models with extensions to correlated serial errors and time-varying parameter models." *Journal of Econometrics* 68, 339-360.

Chib, S. and Greenberg, E. (1998), "Bayesian analysis of the multivariate probit model." *Biometrika* 85, 347-361.

Cochrane, J. H. (1997), "Where is the market going? Uncertain facts and novel theories." *Federal Reserve Bank of Chicago – Economic Perspectives* 21, 3-37.

Cragg, J. G. (1986), "The relationship of dividend payments to the characteristics of the earnings streams of corporations." In *Prices, Competition and Equilibrium*, 191-213, Edited by Peston, M. H. and Quandt, R. E., Barnes and Noble, New York.

Dueker, M. J. (1999), "Conditional heteroscedasticity in qualitative response models of time series: A Gibbs sampling approach to the bank prime rate." *Journal of Business & Economic Statistics* 17, 466-72.

Dueker, M. J. (2005), "Dynamic forecasting of qualitative variables: A Qual VAR Model of U.S. recessions." Forthcoming in *Journal of Business & Economic Statistics.*

Gelfand, A. E. and Smith, A. F. M. (1990), "Sampling-based approaches to calculating marginal densities' *Journal of the American Statistical Association* 85, 398-409.

Geman and Geman (1984), "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 6, 721-741.

Heckman, J. (1974), "Shadow prices, market wages, and labor supply." *Econometrica* 42, 679-694.

Heckman, J. (1979), "Sample selection bias as a specification error." *Econometrica* 47, 153-161.

Huang, River H. C. (2001), "Bayesian analysis of the dividend behavior." *Applied Financial Economics* 11, 333-339.

Huang, River H. C. (2003), "Test of regime-switching CAPM under price limits." *International Review of Economics and Finance* 12, 305-326.

Huang, River H. C. and Shen, C. H. (2002), "Estimation of Taiwan's binary monetary policy reaction function." *Journal of Economic Studies* 29, 222-239.

Kim, B. S. and Maddala, G. S. (1992), "Estimation and specification analysis of models of dividend behavior based on censored panel data." *Empirical Economics* 17, 111-124.

Lintner, J. (1956), "Distribution of incomes of corporations among dividends, retained earnings, and taxes." *American Economic Review* 46, 97-113.

Maddala, G. S. (1983), *Limited–Dependent and Qualitative Variables in Econometrics.* Cambridge University Press, Cambridge.

Scollnik, D. P. M. (1993), "A Bayesian analysis of a simultaneous equations model for insurance rate-making." *Insurance: Mathematics and Economics* 12, 265-286.

Tobin, J. (1958), "Estimation of relationships for limited dependent variables." *Econometrica* 26, 24-36.

Wei, S. X. (2002), "A censored-GARCH model of asset returns with price limits." *Journal of Empirical Finance* 179-223.

Wei, S. X. (1999), "A Bayesian approach to dynamic Tobit models." *Econometric Reviews* 18, 417-439.

Zangari, P. J. (1994), *Maximum likelihood and Bayesian computational procedures for the censored autoregressive model with applications*, Unpublished Ph.D. dissertation, The State University of New Jersey.

Zangari, P. J. and Tsurumi, H. (1996), "A Bayesian analysis of censored autocorrelated data on exports of Japanese passenger cars to the United States." In *Advances in Econometrics: Bayesian Computational Methods and Applications*, 111-143. Edited by Fomby, T. B. and Hill, R. C., JAI Press Inc.