Odejar:

Bayesian Analysis of Sample Selection and Endogenous Switching Regression Models with Random Coefficients Via MCMC Methods

Sonderforschungsbereich 386, Paper 291 (2002)

Online unter: http://epub.ub.uni-muenchen.de/

Bayesian Analysis of Sample Selection and Endogenous Switching Regression Models with Random Coefficients Via MCMC Methods*

Maria Ana Ebron Odejar

Working Paper Sonderforschungsbereich (SFB) 386
Statistical Analysis of Discrete Structures
University of Munich
Statistics and Economics Departments

August 2002

Abstract. This paper develops a Bayesian method for estimating and testing the parameters of the endogenous switching regression model and sample selection models. Random coefficients are incorporated in both the decision and regime regression models to reflect heterogeneity across individual units or clusters and correlation of observations within clusters. The case of tobit type regime regression equations are also considered. A combination of Markov chain Monte Carlo methods, data augmentation and Gibbs sampling is used to facilitate computation of Bayes posterior statistics. A simulation study is conducted to compare estimates from full and reduced blocking schemes and to investigate sensitivity to prior information. The Bayesian methodology is applied to data sets on currency hedging and goods trade, cross-country privatisation, and adoption of soil conservation technology. Estimation and inference results on marginal effects, average decision or selection effect as well as model comparison are presented. The expected decision effect is broken down into average effect of individual's decision on the response variable, decision effect due to random components, and differential effect due to latent correlated random components. Application of the proposed Bayesian MCMC algorithm to real data sets reveal that the normality assumption still holds for most commonly encountered economic data.

Key Words: Sample selection, endogenous switching, Markov chain Monte Carlo methods, data augmentation, Gibbs sampling.

JEL Classification: C11, C12, C13, C15, C31, C33, C34, C35

---

*This research is conducted as a visitor at the University of Munich, Statistics and Economics Departments. I would like to express my gratitude for access to library and computer resources at the University of Munich, Statistics and Economics Departments and financial support from the German National Science Foundation through Sonderforschungsbereich (SFB 386) Statistical Analysis of Discrete Structures. I am especially grateful to Ludwig Fahrmeir of the Statistics Department, and Stephan Klasen and Antonio Garcia Pascual of the Economics Department. I am also thankful to Shang-Jin Wei of NBER, Brookings Institute and Harvard Kennedy School of Government for most of the currency hedging and goods trade data, and to Bernardo Bortolotti and Domenico Siniscalco of FEEM, Milan for the cross-country privatisation data.

1. Introduction

An econometric model that specifies a decision process and the regression models associated with each decision option is the endogenous switching regression model[1] in which the observational units are allocated to a specific regime, depending on the value of the latent decision variable relative to a threshold value. The latent decision variable represents marginal cost and marginal benefit considerations or an individual's expected utility. Sample-selection and disequilibrium models belong to this general class of switching models with the switch determined endogenously (Maddala and Nelson, 1975).Various economic phenomena in labor, migration, health, finance and program evaluation studies can be modelled as an endogenous switching regression model. A firm's decision whether or not to adopt a new technology may be based on productivity gains and cost of adopting the new technology. Self-selection models were used by Gronau (1974), Lewis (1974) and Heckman (1974) to model women's labor force participation decision. Lee and Trost (1978) and Charlier, Melenberg and van Soest (2001) applied it to housing choices of owning or renting. The problem of education and self-selection was analyzed by Willis and Rosen (1979). Mundaca (2001) modelled exchange rate regime switching based on central bank's intervention criteria function as an endogenous switching regression model. A selection model in a study of the quality of hospital care was considered by Gowrisankaran and Town (1999).

1For the stochastic switching regression model, estimation methods are discussed by Quandt and Ramsey (1978) using the moment-generating function technique, Hartley (1978) using the EM algorithm, and Odejar and McNulty (2001) for the EM, data augmentation and Gibbs sampling algorithms.

Recent econometric papers focused on estimating panel data models with unobserved individual specific random effects. Kyriazidou (1997) proposed a two-step estimation method which provides consistent and asymptotically normal estimators for estimating a panel data sample selection model with latent individual specific effects in both the selection and regression equations. In the first step, the unknown coefficients of the selection equation are consistently estimated and in the second stage, the estimates are plugged into the regression equation of interest. In her methodology, the sample selection effect and the unknown coefficients are differenced out from the equation of interest. Barrachina and Engracia (1999) likewise introduced a two-step estimation method for a panel data sample selection model with individual specific effects in both the selection and regression equations. The endogenous switching regression model may also be estimated by the full information maximum likelihood (FIML) method. FIML considers the entire system of equations, and all the parameters are jointly estimated. Estimators obtained by FIML enjoy all the properties of maximum likelihood estimators. They are consistent and asymptotically normally distributed. Most important of all these properties is that the estimators are asymptotically efficient and achieve the Cramer-Rao lower bound. Thus, FIML estimators are most efficient among estimators of the simultaneous equations model which is the endogenous switching regression model in this case. The nonlinear optimization method used to implement FIML is Newton‘s algorithm. Terza (1998) used FIML to estimate an endogenous switching regression model with count data. Compared to two-stage procedures, the FIML method is computationally quite cumbersome to implement, especially with increasing number of regressors. Moreover, it may converge to a local maximum or even to a saddle point. Another criticism of the maximum likelihood method is that it does not provide parameter estimates accurate enough to be useful for small and moderately large samples. It is well known that the t-test yields misleading results when the sample size is small. Nawata and McAleer (2001) demonstrated that this finite sample problem with the t-test is alarming and more severe for binary choice and sample selection models.

Charlier, Melenberg and van Soest (2001) introduced a semiparametric method of analyzing the endogenous switching regression model for panel data. They considered both fixed and random effects models.

In contrast to the maximum likelihood method, the MCMC Bayesian methods are useful and reliable even for finite sample size since convergence results depend only on the number of iterations.

This paper develops a Bayesian method for estimating the parameters of the endogenous switching regression and the standard sample selection models. Heterogeneity across individual units or clusters and association within clusters are accounted for by incorporating random coefficients in both the decision and regression models . Tobit type regime regression equations are also considered. All variables in the model except the dummy variables are assumed to be distributed as log-normal. Markov chain Monte Carlo (MCMC) methods data augmentation and Gibbs sampling are implemented to facilitate computation of posterior estimates. A simulation study is conducted to determine the performance of two MCMC algorithms for varying prior values of the parameters. These Bayesian methods are applied to a currency hedging and bilateral trade study, cross-country privatisation data, and adoption of soil conservation technology study.

The contents of this paper are outlined as follows: Section 2 specifies the endogenous switching regression and the standard sample selection models. Section 3 develops a Bayesian framework for estimating these models. In section 4, MCMC methods and the algorithms for their implementation for the endogenous switching regression and standard sample selection models are discussed. In section 5, simulation results are presented, and in section 6 Bayesian implementation via MCMC methods are applied to real data. Section 7 contains a summary and conclusion.

2. Endogenous Switching Regression Models

In the endogenous switching regression model considered in this paper, the decision process is specified as a linear mixed model

$$d_{ij}^* = z_{ij}'\gamma + w_{ij}'\tau_i + \eta_{ij} \qquad i=1,\ldots,s \quad j=1,\ldots, n_i . \tag{1}$$

Here d_{ij}^* is the latent decision variable, z_{ij} and w_{ij} are the regressors affecting the decision rule, γ are the fixed coefficients, τ_i are the random coefficients assumed to be normal with mean **0** and variance-covariance matrix Σ_τ, η_{ij} is iid normal with mean 0 and variance 1, s is the number of individual units or clusters and n_i is the number of repeated observations over time per individual or number of individual units per cluster. The binary observed decision variable d_{ij} is related to d_{ij}^* through the threshold mechanism

$$d_{ij} = 1 \;,\; d_{ij}^* > 0$$
$$\quad = 0 \;,\; d_{ij}^* \leq 0 \;.$$

If $d_{ij}^* > 0$, that is the marginal benefit or expected utility of belonging to regime 1 is positive then $d_{ij} = 1$, the individual chooses to be in regime 1. Otherwise, the individual decides to belong to regime 2 and $d_{ij} = 0$. The regression models corresponding to each regime are as follows

$$\textit{Regime 1:} \quad y_{ij1} = x_{ij1}'\beta_1 + p_{ij1}'\kappa_{i1} + \varepsilon_{ij1}, \quad \text{if } d_{ij} = 1 \tag{2}$$

$$\textit{Regime 2:} \quad y_{ij0} = x_{ij0}'\beta_0 + p_{ij0}'\kappa_{i0} + \varepsilon_{ij0}, \quad \text{if } d_{ij} = 0\,. \tag{3}$$

The covariances of the errors in the decision and regression models are $\sigma_{\eta\varepsilon_1}$ and $\sigma_{\eta\varepsilon_0}$. The regime errors are assumed to be uncorrelated. The joint model for the latent decision in (1) and regime equations (2) and (3) above may be expressed in matrix form as

$$Y_{ij} = X_{ij}\beta + V_{ij}\kappa_i + \varepsilon_{ij} \tag{4}$$

$$Y_{ij} = \begin{bmatrix} d_{ij}^{\,*} \\ y_{ij1} \\ y_{ij0} \end{bmatrix}, X_{ij} = \begin{bmatrix} z_{ij}' & & \\ & x_{ij1}' & \\ & & x_{ij0}' \end{bmatrix}, \ \beta = \begin{bmatrix} \gamma \\ \beta_1 \\ \beta_0 \end{bmatrix}, \ V_{ij} = \begin{bmatrix} w_{ij}' & & \\ & p_{ij1}' & \\ & & p_{ij0}' \end{bmatrix}$$

$$\kappa_i = \begin{bmatrix} \tau_i \\ \kappa_{i1} \\ \kappa_{i0} \end{bmatrix} \sim \mathrm{N}(\mathbf{0},\ \Sigma_\kappa) \ , \ \Sigma_\kappa = \begin{bmatrix} \Sigma_\tau & & \\ & \Sigma_{\kappa_1} & \\ & & \Sigma_{\kappa_0} \end{bmatrix}$$

$$\varepsilon_{ij} = \begin{bmatrix} \eta_{ij} \\ \varepsilon_{ij1} \\ \varepsilon_{ij0} \end{bmatrix} \sim \mathrm{N}(\mathbf{0},\ \Sigma_\varepsilon) \ , \ \Sigma_\varepsilon = \begin{bmatrix} 1 & \sigma_{\eta\varepsilon_1} & \sigma_{\eta\varepsilon_0} \\ \sigma_{\eta\varepsilon_1} & \sigma_{\varepsilon 1}^{\ 2} & 0 \\ \sigma_{\eta\varepsilon_0} & 0 & \sigma_{\varepsilon 0}^{\ 2} \end{bmatrix}$$

$\boldsymbol{Y}_{ij}$ is the matrix of the response variables including the latent variables, $\boldsymbol{X}_{ij}$ is the matrix of regressors for the fixed coefficients, $\boldsymbol{\beta}$ is the vector of fixed coefficients, $\boldsymbol{V}_{ij}$ is the matrix of covariates for the random coefficients, $\boldsymbol{\kappa}_i$ is the vector of random coefficients and ε_{ij} is the vector of random errors. The conditional distribution of $\boldsymbol{Y}_{ij}$ is multivariate normal $\mathrm{N}(\boldsymbol{M}_{ij}, \boldsymbol{\Sigma}_{y_{ij}})$. Conditional on the fixed coefficients $\boldsymbol{\beta}$ and the random coefficients $\boldsymbol{\kappa}_i$, $\boldsymbol{M}_{ij}$ is $(\boldsymbol{X}_{ij}\boldsymbol{\beta} + \boldsymbol{V}_{ij}\boldsymbol{\kappa}_i)$ and $\boldsymbol{\Sigma}_{y_{ij}}$ is Σ_ε. This will be referred to later in MCMC algorithm1. However, with the conditional distribution of $\boldsymbol{Y}_{ij}$ marginalized over the random coefficients $\boldsymbol{\kappa}_i$ (as in MCMC algorithm 2), $\widetilde{\boldsymbol{M}}_{ij}$ is $\boldsymbol{X}_{ij}\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}_{y_{ij}}$ is $(\boldsymbol{\Sigma}_{\varepsilon} + \boldsymbol{V}_{ij}\boldsymbol{\Sigma}_{\kappa}\boldsymbol{V}_{ij}{}')$.

Tobit Type Endogenous Switching Regression Model

Equations (2) and (3) may also be tobit type models i.e.

$$\begin{aligned} \textit{Regime 1:} \quad y_{ij1} &= y_{ij1}^* = x_{ij1}'\beta_1 + p_{ij1}'\kappa_{i1} + \varepsilon_{ij1}\,,\ \textit{if}\ y_{ij1}^* > 0\ \textit{and}\ d_{ij} = 1 \qquad (5) \\ &= 0 \quad \textit{if}\ y_{ij1}^* \leq 0\ \textit{and}\ d_{ij} = 1 \end{aligned}$$

$$\begin{aligned} \textit{Regime 2:} \quad y_{ij0} &= y_{ij0}^* = x_{ij0}'\beta_0 + p_{ij0}'\kappa_{i0} + \varepsilon_{ij0}\,,\ \textit{if}\ y_{ij0}^* > 0\ \textit{and if}\ d_{ij} = 0 \qquad (6) \\ &= 0 \quad \textit{if}\ y_{ij0}^* \leq 0\ \textit{and}\ d_{ij} = 0\,. \end{aligned}$$

Sample Selection Model

For the sample selection model the response variable is observed only when the latent decision variable $d_{ij}^* > 0$ or consequently when $d_{ij} = 1$. Thus the specification of the sample selection model consists of equation (1) and the following regression model

$$\begin{aligned} y_{ij} &= x_{ij}'\beta + p_{ij}'\kappa_i + \varepsilon_{ij}\,, \quad \textit{if}\ d_{ij} = 1 \qquad (7) \\ &= 0 \qquad\qquad\qquad\quad\ , \quad \textit{if}\ d_{ij} = 0 \end{aligned}$$

3. Bayesian Framework

The likelihood function for the endogenous switching regression model with random coefficients is

$$\mathrm{L}(\beta, \Sigma_{y_i}, \Sigma_{\kappa}, \mathrm{Y}) \propto \prod_{i=1}^{s} \int \prod_{j=1}^{n_i} \boldsymbol{f}\left(\boldsymbol{y}_{ij} \middle| \kappa_i, \beta, \Sigma_{y_i}, \Sigma_{\kappa}\right) \boldsymbol{g}\left(\kappa_i \,|\, \Sigma_{\kappa}\right) \boldsymbol{d}\kappa_i \qquad (8)$$

where

$$\boldsymbol{f}\left(\boldsymbol{y}_{ij}\middle|\kappa_{i},\beta,\boldsymbol{\Sigma}_{y_i},\boldsymbol{\Sigma}_{\kappa}\right)=\left\{\frac{1}{\sigma_{\varepsilon_1}}\phi\left(\frac{\boldsymbol{y}_{ij1}-\boldsymbol{x}_{ij1}'\beta_1-\boldsymbol{p}_{ij1}'\kappa_{i1}}{\sigma_{\varepsilon_1}}\right)\boldsymbol{\Phi}\left[\frac{\boldsymbol{z}_{ij}'\gamma+\boldsymbol{w}_{ij}'\tau_{i}+\frac{\sigma_{\eta\varepsilon_1}}{\sigma_{\varepsilon_1}^{2}}\left(\boldsymbol{y}_{ij1}-\boldsymbol{x}_{ij1}'\beta_1-\boldsymbol{p}_{ij1}'\kappa_{i1}\right)}{\sqrt{1-\frac{\sigma_{\eta\varepsilon_1}^{2}}{\sigma_{\varepsilon_1}^{2}}}}\right]\right\}^{\boldsymbol{d}_{ij}}$$

$$\left\{\frac{1}{\sigma_{\varepsilon_0}}\phi\left(\frac{\boldsymbol{y}_{ij0}-\boldsymbol{x}_{ij0}'\beta_0-\boldsymbol{p}_{ij0}'\kappa_{i0}}{\sigma_{\varepsilon_0}}\right)\boldsymbol{\Phi}\left[\frac{-\boldsymbol{z}_{ij}'\gamma-\boldsymbol{w}_{ij}'\tau_{i}+\frac{\sigma_{\eta\varepsilon_0}}{\sigma_{\varepsilon_0}^{2}}\left(\boldsymbol{y}_{ij0}-\boldsymbol{x}_{ij0}'\beta_0-\boldsymbol{p}_{i0}'\kappa_{i0}\right)}{\sqrt{1-\frac{\sigma_{\eta\varepsilon_0}^{2}}{\sigma_{\varepsilon_0}^{2}}}}\right]\right\}^{1-\boldsymbol{d}_{ij}}.$$

This likelihood is analytically intractable involving high-dimensional integral equal to the number of random coefficients which in this case is s . To ensure a proper posterior density, parameters are modelled with informative priors. It is trivial to obtain prior information from a subset of the sample data when no prior values are available from economic theory or previous research. Moreover, informative priors can easily be adjusted to reflect the degree of certainty or confidence on the priors. The prior distribution for the inverse variance-covariance of the vector of responses and the latent decision variable is $\Sigma_{y_i}^{-1}\sim W\left(\alpha_y, H_y^{-1}\right)$, a Wishart distribution with mean $\alpha_y H_y^{-1}$ and precision matrix H_y^{-1}. Likewise for the hyperparameter $\boldsymbol{\Sigma}_{\kappa}^{-1}$ the prior distribution is $\boldsymbol{\Sigma}_{\kappa}^{-1}\sim \boldsymbol{W}\left(\alpha_{\kappa},\boldsymbol{H}_{\kappa}^{-1}\right)$. The Wishart degrees of freedom for both $\boldsymbol{\Sigma}_{\kappa}^{-1}$ and $\boldsymbol{\Sigma}_{y_i}^{-1}$ has to be small relative to the total sample size to allow the data to dominate the priors. Simulations in this research indicate that a reasonable degrees of freedom is at least 30 for sufficiently accurate estimation. The priors for the random coefficients κ_i and the fixed coefficients β are multivariate normals $\kappa_i \sim N(\mathbf{0}, \Sigma_\kappa)$ *and* $\beta \sim N(\mathbf{0}, \Sigma_\beta)$ respectively.

If we let $g(\beta, \Sigma_{y_i}, \Sigma_\kappa)$, be the joint prior distribution for $(\beta, \Sigma_{y_i}, \Sigma_\kappa)$, the corresponding posterior distribution then is

$$f(\beta, \Sigma_{y_i}, \Sigma_\kappa | Y) = \frac{\prod_{i=1}^{s} \int \prod_{j=1}^{n_i} f\left(y_{ij} \middle| \kappa_i, \beta, \Sigma_{y_i}, \Sigma_\kappa\right) g\left(\kappa_i \middle| \Sigma_\kappa\right) g(\beta, \Sigma_{y_i}, \Sigma_\kappa) d\kappa_i}{\int \prod_{i=1}^{s} \int \prod_{j=1}^{n_i} f\left(y_{ij} \middle| \kappa_i, \beta, \Sigma_{y_i}, \Sigma_\kappa\right) g\left(\kappa_i \middle| \Sigma_\kappa\right) g(\beta, \Sigma_{y_i}, \Sigma_\kappa) d\kappa_i d\beta d\Sigma_{Y_i} d\Sigma_\kappa} \tag{9}$$

Inference on the random coefficients is based on the posterior distribution

$$\boldsymbol{f}(\boldsymbol{b}_i | \boldsymbol{Y}) = \frac{\int \boldsymbol{f}\left(\boldsymbol{y}_{ij} \middle| \kappa_i, \beta, \Sigma_{y_i}, \Sigma_\kappa\right) \boldsymbol{g}\left(\kappa_i | \Sigma_\kappa\right) \boldsymbol{g}(\beta, \Sigma_{y_i}, \Sigma_\kappa)\, \boldsymbol{d}\beta\, \boldsymbol{d}\Sigma_{y_i}\, \boldsymbol{d}\Sigma_\kappa}{\int \boldsymbol{f}\left(\boldsymbol{y}_{ij} \middle| \kappa_i, \beta, \Sigma_{y_i}, \Sigma_\kappa\right) \boldsymbol{g}\left(\kappa_i | \Sigma_\kappa\right) \boldsymbol{g}(\beta, \Sigma_{y_i}, \Sigma_\kappa)\, \boldsymbol{d}\kappa_i\, \boldsymbol{d}\beta\, \boldsymbol{d}\Sigma_{y_i}\, \boldsymbol{d}\Sigma_\kappa}. \tag{10}$$

These posterior distributions are too complicated to evaluate analytically. However, by using MCMC methods data augmentation and Gibbs sampling, this posterior distribution can be sampled indirectly by generating a sample of parameter values from the conditional distributions which are relatively simpler in form than the joint posterior distribution of interest. Posterior Bayes estimates are then obtained from the generated samples.

4. Markov Chain Monte Carlo Methods

Markov chain Monte Carlo (MCMC) methods facilitate Bayesian estimation of the endogenous switching regression parameters. MCMC methods aim to summarize the features of a distribution by sampling indirectly from the distribution of interest. These methods are most valuable for complicated distributions such as high dimensional joint distributions which are analytically infeasible to evaluate. MCMC methods construct a

Markov chain $\theta_k^{(1)}, \theta_k^{(2)}, \ldots, \theta_k^{(m)}, \ldots$ with equilibrium distribution identical to the desired joint posterior distribution. Ergodic averaging of the Markov chain $\theta_k^{(m)}$ or some function $h(\theta_k^{(m)})$ provides consistent estimators of the parameters θ or a function $h(\theta)$. A form of MCMC is Gibbs (Geman and Geman 1984) sampling where Markov chains are generated from full conditional distributions $f(\theta_k \mid \theta_1,\ldots,\theta_{k-1}, \theta_{k+1},\ldots,\theta_r)$, $k=1,\ldots,r$. The resulting iterations from

(1) Generating $\theta_1^{(m+1)}$ from $f(\theta_1^{(m+1)} \mid \theta_2^{(m)},\ldots,\theta_r^{(m)})$.

(2) Generating $\theta_2^{(m+1)}$ from $f(\theta_2^{(m+1)} \mid \theta_1^{(m+1)}, \theta_3^{(m)},\ldots,\theta_r^{(m)})$.

.

.

.

(r) Generating $\theta_r^{(m+1)}$ from $f(\theta_r^{(m+1)} \mid \theta_1^{(m+1)},\ldots,\theta_{r-1}^{(m+1)})$

provide a Markov chain with transition probability from $\theta^{(m)}$ to $\theta^{(m+1)}$ given by the product of the above r full conditional probabilities. Under regulatory conditions (Tierney 1991), as the number of iterations m approaches infinity $(\theta_1^{(m)},\ldots, \theta_r^{(m)})$ converges in distribution to $(\theta_1,\ldots, \theta_r)$ and likewise $\theta_k^{(m)}$ converges in distribution to θ_k. After equilibrium is reached at iteration a, sample values are averaged to provide consistent estimates of the parameters or their function,

$$\hat{E}[h(\theta_k)] = \frac{\sum_{m=a+1}^{t} h(\theta_k)^{(m)}}{t-a} \quad . \tag{11}$$

The marginal posterior distribution is estimated as

$$f(\theta_k) \approx \frac{\sum_{m=a+1}^{t} f\left[\theta_k \middle| \left(\theta_1^{(m)},\ldots\theta_{k-1}^{(m)},\theta_{k+1}^{(m)}\ldots\theta_r^{(m)}\right)\right]}{t-a} \tag{12}$$

and the estimate of the conditional predictive ordinate (CPO) is

$$\int f\left(Y^f \middle| \theta, X^f, V^f\right) g\left(\theta \middle| Y, X, V\right) d\theta \approx \frac{\sum_{m=a+1}^{t} f\left(Y^f \middle| \theta^{(m)}, X^f, V^f\right)}{t-a} \quad . \tag{13}$$

In contrast to the maximum likelihood method and the bootstrap resampling method, the MCMC Bayesian methods are useful even for finite sample size since convergence results depend only on the number of iterations.

In data augmentation (Tanner and Wong 1987), a form of Gibbs sampling, there are only two blocks $f(M_Y| \ \theta)$ and $f(\theta \mid M_Y)$ corresponding to an imputation step and a posterior step where M_Y is the vector of latent variables that augment the original data and simplifies analysis of complicated models.

The key to analyzing the endogenous switching regression model is to apply data augmentation to generate the missing variables $M_y = [d_{ij}^*, y_{ij1}^{\ m}, y_{ij0}^{\ m}]$ where d_{ij}^* is the latent decision variable and $y_{ij1}^{\ m}$ and $y_{ij0}^{\ m}$ are the potential response variables if $d_{ij} =0$ and if $d_{ij} =1$ respectively. With the original data plus the latent decision variable and potential variables known, the data is complete and evaluation of the likelihood and the joint posterior distribution is greatly simplified. Hence, the likelihood in (8) simplifies to that of the multivariate linear model

$$\mathbf{L(\theta,Y)} \propto \exp\left[-\frac{1}{2} tr \sum_{i=1}^{s} \Sigma_{Y_i}^{-1} \left(Y_i - X_i\beta - V_i\kappa_i\right)\left(Y_i - X_i\beta - V_i\kappa_i\right)'\right] \tag{14}$$

for balanced clusters, where $\Sigma_{y_i} = I_{n_i} \otimes \Sigma_\varepsilon$

and

$$\mathbf{L(\theta\ ,Y)} \quad \propto \exp\left[-\frac{1}{2} tr \sum_{i=1}^{s}\sum_{j=1}^{n_i} \Sigma_{Y_{ij}}^{-1} \left(Y_{ij} - X_{ij}\beta - V_{ij}\kappa_i\right)\left(Y_{ij} - X_{ij}\beta - V_{ij}\kappa_i\right)'\right] \tag{15}$$

for clusters with unequal sample size.

When $d_{ij}=1$, that is the decision is to select the first option, d_{ij}^* is generated from the conditional distribution $[d_{ij}^* \mid \theta] \sim TN_{(0,\infty)}\, [\, (z_{ij}'\gamma + w_{ij}'\tau_i)\, ,\, 1\,]$ which is a truncated normal distribution with support $(0,\infty)$. Using the inversion method (Devroye 1986), samples from the truncated normal are therefore generated from $(z_{ij}'\gamma + w_{ij}'\tau_i) + \Phi^{-1}(1-\Phi(z_{ij}'\gamma + w_{ij}'\tau_i) + U\Phi(z_{ij}'\gamma + w_{ij}'\tau_i))$ where U is the standard uniform distribution. Rejection method and Geweke's (1991) method can also be used to sample from the truncated normal. The latent potential variable y_{ij0}^m, is then generated from the untruncated normal distribution

$$[y_{ij0}^{\,m} \mid d_{ij}^*, \theta] \sim N[x_{ij1}'\beta_0 + p_{ij1}'\kappa_{io} + \sigma_{\eta\varepsilon_0}(d_{ij}^* - z_{ij}'\gamma - w_{ij}'\tau_i)\, ,\, \sigma_{\varepsilon_0}^{\;2} - \sigma_{\eta\varepsilon_0}^{\;2}\,]$$

If the decision is to choose option 2, that is $d_{ij}=0$, the latent decision variable d_{ij}^* is then generated from the truncated normal $[d_{ij}^* \mid \theta] \sim TN_{(-\infty,\,0)}\, [\, (z_{ij}'\gamma + w_{ij}'\tau_i)\, ,\, 1\,]$ with support $(-\infty\, ,\, 0)$. Using the inversion method samples are drawn from $(z_{ij}'\gamma + w_{ij}'\tau_i) + \Phi^{-1}(\, U(1-\Phi(z_{ij}'\gamma + w_{ij}'\tau_i)))$. The latent potential variable y_{ij1}^m, is generated from the conditional distribution $[y_{ij1}^{\,m} \mid d_{ij}^*\, ,\, \theta] \sim N[x_{ij0}'\beta_1 + p_{ij0}'\kappa_{i1} + \sigma_{\eta\varepsilon_1}(d_{ij}^* - z_{ij}'\gamma - w_{ij}'\tau_i)\, ,\, \sigma_{\varepsilon_1}^{\;2} - \sigma_{\eta\varepsilon_1}^{\;2}\,]$.

Algorithm 1

The complete Gibbs sampling algorithm for a two regime endogenous switching regression model in (1)-(3) with initial values $\gamma^{(0)}$, $\tau_i^{(0)}$, $\beta_1^{(0)}$, $\beta_0^{(0)}$, $\kappa_{i1}^{(0)}$, $\kappa_{i0}^{(0)}$, $\sigma_{\eta\varepsilon_0}^{\;(0)}$, $\sigma_{\eta\varepsilon_1}^{\;(0)}$, $\sigma_{\varepsilon_0}^{\;2(0)}$,$\sigma_{\varepsilon_1}^{\;2(0)}$ proceeds as follows:

<u>Imputation Step:</u>

Generate $\boldsymbol{M}_y^{(m)}$ from $[\boldsymbol{M}_y^{(m)} \mid \theta^{(m-1)}]$ as follows:

1a. Generate $d_{ij}^{*(m)}$ from

$[d_{ij}^{*} \mid \theta^{(m-1)}] \sim TN_{(0,\infty)} [(z_{ij}`\gamma + w_{ij}`\tau_i) , 1]$ with support $(0,\infty)$ if $d_{ij}=1$.

1b. Generate $y_{ij0}^{(m)}$ from

$$[y_{ij0}^{m} \mid d_{ij}^{*\,(m)}, \theta^{(m-1)}] \sim N[x_{ij1}'\beta_0 + p_{ij1}'\kappa_{io} + \sigma_{\eta\varepsilon_0} (d_{ij}^{*}-z_{ij}`\gamma-w_{ij}`\tau_i) , \sigma_{\varepsilon_0}^{2} - \sigma_{\eta\varepsilon_0}^{2}].$$

2a. Generate $d_{ij}^{*(m)}$ from

$[d_{ij}^{*} \mid \theta^{(m-1)}] \sim TN_{(-\infty,\,0)} [(z_{ij}`\gamma + w_{ij}`\tau_i) , 1]$ with support $(-\infty, 0)$ if $d_{ij}=0$.

2b. Generate $y_{ij1}^{m\ (m)}$ from

$$[y_{ij1}^{m} \mid d_{ij}^{*\,(m)}, \theta^{(m-1)}] \sim N[x_{ij0}'\beta_1 + p_{ij0}'\kappa_{i1} + \sigma_{\eta\varepsilon_1} (d_{ij}^{*}-z_{ij}`\gamma-w_{ij}`\tau_i) , \sigma_{\varepsilon_1}^{2} - \sigma_{\eta\varepsilon_1}^{2}].$$

<u>Posterior Step:</u>

Generate $\theta^{(m)}$ from $[\theta \mid \boldsymbol{M}_y^{(m)}]$ as follows:

3. Generate $\boldsymbol{\Sigma}_{\kappa}^{-1(\boldsymbol{m})}$ from the Wishart

$$\left[\boldsymbol{\Sigma}_{\kappa}^{-1} \middle| \boldsymbol{\kappa}_i^{(\boldsymbol{m}-1)}, \boldsymbol{M}_y^{(\boldsymbol{m})}\right] \sim \boldsymbol{W}\left\{(\alpha_{\kappa} + \boldsymbol{s}), \left[\boldsymbol{H}_{\kappa} + \sum_{\boldsymbol{i}=1}^{\boldsymbol{s}} \boldsymbol{\kappa}_i \boldsymbol{\kappa}_i'\right]^{-1}\right\}$$

by first sampling $W^{*}=TT'$ from a standard Wishart $W((\alpha_{\kappa} + s),I)$ with parameters $(\alpha_{\kappa}+ s)$ and I using Bartlett's (1933) decomposition method described in Ripley (1987). $\boldsymbol{\Sigma}_{\kappa}^{-1(\boldsymbol{m})}$ is then $LW^{*}L'$ where the Choleski decomposition of the precision parameter $\left[H_{\kappa} + \sum_{i=1}^{s} \kappa_i \kappa_i'\right]^{-1}$ is LL'.

4. Generate $\boldsymbol{\Sigma}_{y_i}^{-1(\boldsymbol{m})}$ from the Wishart distribution

$$\left[\boldsymbol{\Sigma}_{y_i}^{-1} \middle| \beta^{(\boldsymbol{m}-1)}, \boldsymbol{\kappa}_i^{(\boldsymbol{m}-1)}, \boldsymbol{\Sigma}_{\kappa}^{-1(\boldsymbol{m})}, \boldsymbol{M}_y^{(\boldsymbol{m})}\right] \sim \boldsymbol{W}\left\{(\alpha_y + \boldsymbol{s}), \left[\boldsymbol{H}_y + \sum_{\boldsymbol{i}=1}^{\boldsymbol{s}} (\boldsymbol{Y}_i - \boldsymbol{X}_i\beta - \boldsymbol{P}_i\boldsymbol{\kappa}_i)(\boldsymbol{Y}_i - \boldsymbol{X}_i\beta - \boldsymbol{P}_i\boldsymbol{\kappa}_i)'\right]^{-1}\right\}$$

using the method above.

5. Generate $\kappa_i^{(m)}$ from $[\boldsymbol{\kappa}_i \,|\, \boldsymbol{\Sigma}_{\kappa}^{-1(\boldsymbol{m})}, \beta^{(\boldsymbol{m}-1)}, \boldsymbol{\Sigma}_{y_i}^{-1(\boldsymbol{m})}, \boldsymbol{M}_y^{(\boldsymbol{m})}] \sim \boldsymbol{N}(\boldsymbol{M}_{\kappa}, \boldsymbol{V}_{\kappa})$ where

$$\boldsymbol{V}_{\kappa} = \left(\boldsymbol{\Sigma}_{\kappa}^{-1} + \boldsymbol{P}_i'\boldsymbol{\Sigma}_{y_i}^{-1}\boldsymbol{P}_i\right)^{-1} \quad \text{and} \quad \boldsymbol{M}_{\kappa} = \boldsymbol{V}_{\kappa}\left[\ \boldsymbol{P}_i'\boldsymbol{\Sigma}_{y_i}^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i\beta\right)\right]$$

6. Generate $\beta^{(m)}$ from $[\beta \,|\, \kappa_i^{(m)}, \boldsymbol{\Sigma}_{\kappa}^{-1(m)}, \boldsymbol{\Sigma}_{y_i}^{-1(m)}, \boldsymbol{M}_y^{(m)}] \sim \boldsymbol{N}\left(\boldsymbol{M}_{\beta}, \boldsymbol{V}_{\beta}\right)$

where $\boldsymbol{V}_{\beta} = \left(\boldsymbol{\Sigma}_{\beta}^{-1} + \sum_{i=1}^{s}\boldsymbol{X}_i'\boldsymbol{\Sigma}_{y_i}^{-1}\boldsymbol{X}_i\right)^{-1}$ and $\boldsymbol{M}_{\beta} = \boldsymbol{V}_{\beta}\left[\boldsymbol{\Sigma}_{\beta}^{-1}\boldsymbol{A} + \sum_{i=1}^{s}\boldsymbol{X}_i'\boldsymbol{\Sigma}_{y_i}^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{P}_i\kappa_i\right)\right]$.

In generating the variance-covariance matrices in (3) and (4), the simulations ensure that these matrices are positive definite.

Gibbs sampling is not only a powerful technique but it is also flexible. For the sample selection model, there are only two rows in the matrices in (4) with the corresponding columns also deleted. Thus the above algorithm is modified by skipping (1b) with all other steps unchanged.

For a tobit type endogenous switching regression model, the above algorithm can be easily modified by two additional steps (1c) and (2c) to generate additional latent variables corresponding to the censored values or zero's . In (1c) y_{ij1}^{*m} is generated from the truncated normal

$[y_{ij1}^{*m} \mid d_{ij}^*, \theta^{(m-1)}] \sim TN_{(-\infty,\,0)}\,[x_{ij1}'\beta_1 + p_{ij1}'\kappa_{i1} + \sigma_{\eta\varepsilon_1}(d_{ij}^* - z_{ij}'\gamma - w_{ij}'\tau_i),\ \sigma_{\varepsilon_1}^2 - \sigma_{\eta\varepsilon_1}^2]$

with support $(-\infty, 0)$ for $y_{ij1}=0$. In (2c) y_{ij0}^{*m} is generated from the truncated normal

$[y_{ij0}^{*m} \mid d_{ij}^*, \theta^{(m-1)}] \sim TN_{(-\infty,\,0)}\,[x_{ij0}'\beta_0 + p_{ij0}'\kappa_{i0} + \sigma_{\eta\varepsilon_0}(d_{ij}^* - z_{ij}'\gamma - w_{ij}'\tau_i),\ \sigma_{\varepsilon_0}^2 - \sigma_{\eta\varepsilon_0}^2]$

with support $(-\infty, 0)$ for $y_{ij0}=0$.

If the random coefficients are not included in both the decision and regression equations, the algorithm excludes (3) and (5), and in all the conditional distributions in the other steps, the terms corresponding to the random coefficients are not included.

Parameters are considered significant at the 5% significance level if the interval between the 2.5th and 97.5th percentiles of the MCMC samples from the posterior distribution exclude zero.

For individual units that selected regime one, the actual decision effect for individual j in cluster i may be measured by comparing the actual response for regime one to the expected potential response for regime two,

$$y_{ij1} - E(y_{ij0} / d_{ij}=1) = y_{ij1} - \left[x_{ij1}{}' \beta_0 + p_{ij1}{}' \kappa_{i0} + \sigma_{\eta\varepsilon_0} \lambda_1 \right] \tag{16}$$

where $\lambda_1 = \phi(z_{ij}'\gamma + w_{ij}'\tau_i) \big/ \Phi(z_{ij}'\gamma + w_{ij}'\tau_i)$.

On the other hand, the average or expected decision effect for individual j in cluster i for those that decided to be in regime 1 is

$$E(y_{ij1} / d_{ij}=1) - E(y_{ij0} / d_{ij}=1) \tag{17}$$

$$= \left[x_{ij1}{}' \beta_1 + p_{ij1}{}' \kappa_{i1} + \sigma_{\eta\varepsilon_1} \lambda_1 \right] - \left[x_{ij1}{}' \beta_0 + p_{ij1}{}' \kappa_{i0} + \sigma_{\eta\varepsilon_0} \lambda_1 \right]$$

$$= x_{ij1}{}' (\beta_1 - \beta_0) + p_{ij1}{}' (\kappa_{i1} - \kappa_{i0}) + (\sigma_{\eta\varepsilon_1} - \sigma_{\eta\varepsilon_0}) \lambda_1 .$$

It is the difference between expected response from choosing to be in regime 1 and the expected potential response from opting to be in regime 2. Thus, the decision effect also referred as counterfactual or conditional effect may be decomposed into three parts, the average effect of the individual's decision on the response variable, decision effect due to random components and the differential effect of the decision on the response variable due to the unobserved correlated random components. Significance of each of these terms can be tested from the MCMC samples of the posterior distributions. For program evaluation studies, the total actual effect of the program may be evaluated by summing

(16) across individuals and clusters, while the total average effect of the program maybe evaluated by summing (17) across individuals and clusters. The average program effect may also be evaluated at the mean values of the regressors. This saves a lot of computer time than evaluating the average program effect at each individual observations. If any of the three components in (17) is significant, we can infer that the average decision effect is also significant without actually evaluating (17) for each individual unit and cluster.

The marginal effect of x_{ij1h} on the response variable y_{ij1} is

$$\partial E(y_{ij1} / d_{ij} = 1) \Big/ \partial x_{ij1h} = \beta_h + \gamma_h \sigma_{\eta\varepsilon_1} \delta_1 \tag{18}$$

where $\lambda_1 = \phi(z_{ij}'\gamma + w_{ij}'\tau_i) \Big/ \Phi(z_{ij}'\gamma + w_{ij}'\tau_i)$ and $\delta_1 = \lambda_1^{\;2} + \lambda_1 (z_{ij}'\gamma + w_{ij}'\tau_i)$.

Again the estimates of the marginal effects and their significance can be tested from the MCMC samples of the posterior distributions. Actual and average decision effects and marginal effects for observations that belong to regime 2 are similarly evaluated.

Algorithm 2

A reduced blocking scheme is also implemented in an attempt to accelerate convergence and improve mixing properties of the Markov chains by grouping together elements of the parameter vector that are highly correlated (Liu, Wong and Kong, 1994) and using available reduced conditional distributions (Gelfand and Smith, 1990). In algorithm 2, the imputation steps are unchanged . However, in the posterior step, the fixed coefficients β and the inverse-variance covariance $\Sigma_{y_i}^{-1}$ are generated in one block and the random coefficients κ_i and their inverse variance covariance matrix $\boldsymbol{\Sigma}_{\kappa}^{-1}$ are drawn in another block. The reduced blocking algorithm 2 procedure is then

<u>Imputation Step:</u>

Generate $\boldsymbol{M}_y^{(m)}$ from $[\boldsymbol{M}_y \mid \theta^{(m-1)}]$.

<u>Posterior Step:</u>

Generate $\boldsymbol{\theta}^{(m)}$ from $[\theta \mid \boldsymbol{M}_y^{(m)}]$.

1. Generate $\beta^{(m)}$ and $\boldsymbol{\Sigma}_{y_i}^{-1\,(m)}$ from $[(\beta, \boldsymbol{\Sigma}_{y_i}^{-1}) \mid \boldsymbol{M}_y^{(m)}]$

1a.Generate $\Sigma_{y_i}^{-1(m)}$ from the Wishart distribution

$$\left[\boldsymbol{\Sigma}_{y_i}^{-1} \mid \beta^{(m-1)}, \boldsymbol{M}_y^{(m)}\right] \sim \boldsymbol{W}\left\{(\alpha_y + \boldsymbol{s}), \left[\boldsymbol{H}_y + \sum_{i=1}^{s} (\boldsymbol{Y}_i - \boldsymbol{X}_i\beta)(\boldsymbol{Y}_i - \boldsymbol{X}_i\beta)'\right]^{-1}\right\}$$

where $\boldsymbol{\Sigma}_{y_i} = (\boldsymbol{V}_i \boldsymbol{\Sigma}_{\kappa} \boldsymbol{V}_i' + \boldsymbol{I}_{n_i} \otimes \boldsymbol{\Sigma}_{\varepsilon})$.

1b. Generate $\beta^{(m)}$ from $[\beta \mid \boldsymbol{\Sigma}_{y_i}^{-1(m)}, \boldsymbol{M}_y^{(m)}] \sim \boldsymbol{N}(\boldsymbol{M}_{\beta}, \boldsymbol{V}_{\beta})$

where $\boldsymbol{V}_{\beta} = \left(\boldsymbol{\Sigma}_{\beta}^{-1} + \sum_{i=1}^{s} \boldsymbol{X}_i' \boldsymbol{\Sigma}_{y_i}^{-1} \boldsymbol{X}_i\right)^{-1}$ and $\boldsymbol{M}_{\beta} = \boldsymbol{V}_{\beta}\left[\boldsymbol{\Sigma}_{\beta}^{-1}\boldsymbol{A} + \sum_{i=1}^{s} \boldsymbol{X}_i' \boldsymbol{\Sigma}_{y_i}^{-1} \boldsymbol{Y}_i\right]$.

2.Generate $\boldsymbol{\Sigma}_{\kappa}^{-1(m)}$ and $\kappa_i^{(m)}$ from $[(\boldsymbol{\Sigma}_{\kappa}^{-1(m)}, \kappa_i^{(m)}) \mid (\beta, \boldsymbol{\Sigma}_{y_i}^{-1})^{(m)}, \boldsymbol{M}_y^{(m)}]$

2a. Generate $\boldsymbol{\Sigma}_{\kappa}^{-1(m)}$ from the Wishart

$$\left[\boldsymbol{\Sigma}_{\kappa}^{-1} \mid \kappa_i^{(m-1)}, (\beta, \boldsymbol{\Sigma}_{y_i}^{-1})^{(m)}, \boldsymbol{M}_y^{(m)}\right] \sim \boldsymbol{W}\left\{(\alpha_{\kappa} + \boldsymbol{s}), \left[\boldsymbol{H}_{\kappa} + \sum_{i=1}^{s} \kappa_i \kappa_i'\right]^{-1}\right\}.$$

2b. Generate $\kappa_i^{(m)}$ from

$$[\kappa_i \mid \boldsymbol{\Sigma}_{\kappa}^{-1(m)}, (\beta, \boldsymbol{\Sigma}_{y_i}^{-1})^{m}, \boldsymbol{M}_y^{(m)}] \sim \boldsymbol{N}(\boldsymbol{M}_{\kappa}, \boldsymbol{V}_{\kappa})$$

where $\boldsymbol{V}_{\kappa} = \left(\boldsymbol{\Sigma}_{\kappa}^{-1} + \boldsymbol{P}_i' \boldsymbol{\Sigma}_{y_i}^{-1} \boldsymbol{P}_i\right)^{-1}$ and $\boldsymbol{M}_{\kappa} = \boldsymbol{V}_{\kappa}\left[\ \boldsymbol{P}_i' \boldsymbol{\Sigma}_{y_i}^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{X}_i \beta\right)\right]$.

5. Simulation Study

To compare parameter estimates for the algorithms with full and reduced MCMC blocking scheme and to assess robustness to prior information, 100 data sets were generated each with 50 clusters and 20 observations per cluster. The endogenous switching regression model considered in the simulations is

$$d_{ij}^* = z_{ij}'\gamma + w_{ij}'\tau_i + \eta_{ij} \qquad i=1,\ldots,50 \quad j=1,\ldots,20. \tag{19}$$

Regime 1: $y_{ij1} = x_{ij1}'\beta_1 + p_{ij1}'\kappa_{i1} + \varepsilon_{ij1}$, *if* $d_{ij} = 1$

Regime 2: $y_{ij0} = x_{ij0}'\beta_0 + p_{ij0}'\kappa_{i0} + \varepsilon_{ij0}$, *if* $d_{ij} = 0$

where $\gamma' = [1, -0.004, 0.3]$, $\beta_1' = [14, -0.3, 0.53]$, $\beta_0' = [-1, 0.7, -0.3]$,

$$\tau_i = \begin{bmatrix} \tau_{i0} \\ \tau_{i1} \end{bmatrix} \sim \mathrm{N}(\mathbf{0}, \Sigma_\tau) \ , \ \Sigma_\tau = \begin{bmatrix} 0.1 & \\ & 0.07 \end{bmatrix}$$

$$\kappa_{i1} = \begin{bmatrix} \kappa_{i10} \\ \kappa_{i11} \end{bmatrix} \sim \mathrm{N}(\mathbf{0}, \Sigma_{\kappa_1}) \ , \ \Sigma_{\kappa_1} = \begin{bmatrix} 0.1 & \\ & 0.06 \end{bmatrix}$$

$$\kappa_{i0} = \begin{bmatrix} \kappa_{i00} \\ \kappa_{i01} \end{bmatrix} \sim \mathrm{N}(\mathbf{0}, \Sigma_{\kappa_0}) \ , \ \Sigma_{\kappa_0} = \begin{bmatrix} 0.1 & \\ & 0.05 \end{bmatrix}$$

$$\varepsilon_{ij} = \begin{bmatrix} \eta_{ij} \\ \varepsilon_{ij1} \\ \varepsilon_{ij0} \end{bmatrix} \sim \mathrm{N}(\mathbf{0}, \Sigma_\varepsilon) \ , \ \Sigma_\varepsilon = \begin{bmatrix} 1 & 0.4 & -0.2 \\ 0.4 & 9 & 0 \\ -0.2 & 0 & 2 \end{bmatrix}.$$

For each data set, each MCMC algorithm is implemented with a burn in or warm up period of *a=2000* iterations and monitoring period of *t=5000* iterations. Estimates of the posterior mean $\widehat{\theta}$, standard deviation $\hat{s}_\theta$, and the 2.5% and 97.5% quantiles of the marginal posterior distributions were computed for each parameter and for each data set.

Tables 1-5 present the true parameter values θ, the average parameter estimate $\bar{\hat{\theta}}$ over 100 data sets, the standard deviation of the parameter estimates $s_{\hat{\theta}}$ over 100 data sets, the average estimate of standard deviation $\bar{\hat{s}}_{\theta}$ over 100 data sets, mean absolute deviation from the true parameter values, and the actual coverage of the nominal 95% Bayes interval from the 2.5th to 97.5th percentiles for varying priors. Three different prior combinations were used $\beta_p = 0.9\beta$, Hy = $0.9\Sigma_{y_i}$ Hκ= $0.9\Sigma_{\kappa}$, $\beta_p = 0.7\beta$, Hy = $0.7\Sigma_{y_i}$ Hκ= $0.7\Sigma_{\kappa}$ and $\beta_p = 0.5\beta$, Hy = $0.5\Sigma_{y_i}$ Hκ= $0.5\Sigma_{\kappa}$. Prior values are specified as certain fractions of the true parameter values. Simulation results show that there is not much difference in posterior statistics obtained from the full and reduced blocking algorithms. This conforms with the results of Gelfand and Smith (1990) that the gain in efficiency from using substitution sampling or data augmentation relative to Gibbs sampling is only likely to be of consequence when the number of reduced conditionals is a relatively large fraction of the total number of conditionals involved in a cycle. Sample paths for the fixed effects are very stable for both algorithms. However, the reduced blocking algorithm 2 has more stable sample paths for the inverse variance-covariance matrix of the error components $\Sigma_{y_i}{}^{-1}$ and the random coefficients $\Sigma_{\kappa}{}^{-1}$ than full blocking algorithm 1. That is, algorithm 2 rarely gets trapped in an absorbing state which is extremely large values for $\Sigma_{\kappa}{}^{-1}$ and extremely small values for $\Sigma_{y_i}{}^{-1}$. These can be remedied by either reinitialising these chains or setting extremely large generated parameter values to a reasonable upper bound like a function of the maximum response values and extremely small generated parameter values to a reasonable lower bound. In the simulation results presented here, the absorbing states are not observed. However, for some parameter values I have also tried (but results are not presented here) from which the data sets were generated in the simulations, the absorbing states sometimes occur. For the privatisation and hedgerow adoption data sets these absorbing states were observed but not for the currency hedging and goods trade data. The occurrence of an absorbing state for Gibbs Markov chain for the variance-covariance matrix of the random effects Σ_{κ} was also observed by Zeger and Karim (1991) for GLM models with random effects. Moreover, the reduced blocking algorithm 2 provided model estimates with better fit in terms of the sum of log conditional posterior ordinate as evidenced by real data analysis for the privatisation study. For both algorithms, estimation and inference results are fairly robust to prior information and parameter estimates are quite accurate especially for the fixed parameters. The standard deviation estimates $\bar{\hat{s}}_{\theta}$ are approximately unbiased providing consistent inferences. The actual coverage probabilities of the nominal 95% Bayes posterior intervals is 100% for all parameters using any prior.

6. Real Data Applications

The proposed Bayesian MCMC method for parameter estimation and hypothesis testing are utilized for analysing real data on currency hedging and goods trade, cross-country privatisation study, and adoption of soil conservation technology. For the real data sets, the Bayesian method is implemented using MCMC method algorithm 2.

Currency Hedging and Goods Trade

The currency hedging and goods trade data of 45 bilateral trade partners for 1970, 1980, 1990, and 1992 is analysed using Bayesian MCMC method for a two regime endogenous switching regression with fixed effects, random intercepts and known threshold value. The decision model which specifies which country pairs are likely to have developed hedging instruments based on the value of the latent variable indicating virtual trade or trade level in the absence of any hedging instrument is

$$d_{ij}^* = \tau_{0i} + \sum_{k=1}^{4} \gamma_k z_{ijk} + \eta_{ij}, \quad i=1,\ldots,45 \;,\; j=1,\ldots,4 \tag{20}$$

where τ_{0i} is the random intercept specific to the ith bilateral trade partners, the regressors z_{ijk} for the fixed coefficients are distance between the economic centers of the two countries (dist), real exchange rate volatility as measured by the standard deviation of the first difference of the natural logarithm of monthly exchange rates during the year (xrvol), the product of two countries per capita GNPs (gnpc) and a dummy variable for country pairs with common language or some historic/colonial ties (comlink). It is assumed that a hedging market will emerge if the level of potential trade d_{ij}^* exceeds 10.5. The bilateral trade level corresponding to this regime is

$$y_{ij1} = \kappa_{i10} + \sum_{k=1}^{4} \beta_{k1} x_{ijk1} + \varepsilon_{ij1} \quad i=1,\ldots,45 \;,\; j=1,\ldots,4, \tag{21}$$

otherwise, the trade level is

$$y_{ij0} = \kappa_{i00} + \sum_{k=1}^{4} \beta_{k0} x_{ijk0} + \varepsilon_{ij0}\,. \tag{22}$$

The regression equations are basically of the same form as the decision equation except that the dependent variable is total trade between the country pairs and the regressors are dist, xrvol, comlink and gnp which is the product of country pairs GNPs. For both the decision and regime equations, all variables except the dummy variables are in natural logarithm. The choice of variables and gravity specification are from Wei (1999). An explanation proposed for the difficulty in identifying large negative effects of exchange rate volatility on trade is the hedging hypothesis which states that the availability of hedging instruments reduces the effect of exchange rate volatility. If the hedging hypothesis holds, the volatility elasticity of trade is reduced possibly to zero for bilateral trade partners that have access to hedging instruments. From table 6, it can be inferred that the hedging hypothesis can be rejected at the 5% significance level. For country pairs with large trade potential and most likely to develop currency hedging instruments, exchange rate volatility deters goods trade enormously. On the other hand, for bilateral trade partners with small trade potential, stochastic volatility has small positive effect on goods trade. Results presented in table 8 show that the difference in stochastic volatility effects on trade is statistically significant at the 5% level. Since the first and third components of the average decision effect in (17) are statistically significant at the 5% level, we can also conclude that the average effect of hedging on trade is significant.

The quantile-quantile (Q-Q) plot in figure 3 shows that the normality assumption of errors is satisfied. The residual plot does not deviate far from the normal line.

Cross-Country Privatisation Study

The Bayesian methodology for sample selection with fixed effects, random coefficients and known threshold value is applied to a cross-country privatisation study of Bortolotti, Fantini and Siniscalco (2001). The data consists of observations from 23 countries for 16 annual periods. It is assumed that governments privatise when their expected utility $d_{ij}*$ exceeds an unobservable threshold which is assumed to be zero in this study. Thus the decision model is specified as

$$d_{ij}{}^* = \tau_{0i} + \sum_{k=1}^{6} \gamma_k z_{ijk} + \eta_{ij}, \quad i=1,\dots,23 \ , \ j=1,\dots,16 \tag{23}$$

where τ_{0i} is the country-specific random intercept, the regressors z_{ijk} for the fixed coefficients are lagged capitalisation (capl), lagged debt to gdp ratio (debtl), per capita gdp (gdp), lagged turnover ratio (turno), and dummy variables for a non-democratic form of government (nondem) and for a right-wing government (right). The regression equation is

$$y_{ij} = \kappa_{0i} + \sum_{k=1}^{4} \beta_k x_{ijk} + \sum_{k=1}^{4} \sum_{k'=1}^{4} \beta_{kk'} x_{ijk} x_{ijk'} + \varepsilon_{ij}, \quad i=1,\dots,23 \ , \ j=1,\dots,16 \tag{24}$$

where the dependent variable rev/gdp is the total gross revenues from privatisation sales scaled by gdp, κ_{0i} is the country-specific random intercept, the regressors x_{ijk} for the fixed coefficients are capl, debtl, gdp and turno. All variables except the dummy variables for both the decision and state regression models are in natural logarithm. From table 9 it can be inferred that privatisation apparently coincides with financial market development as well as economic development and high foreign debt level. Coefficients of lagged capitalisation, lagged debt to gdp ratio, lagged turnover ratio and per capita gdp are all positive and significant at the 5% level. Moreover, right-wing governments are more likely to resort to privatisation. Table 9 shows that rev/gdp varies with capl , debtl and turno in a quadratic manner. Interactions among stock market liquidity, economic development and debt level are manifested as interactions between capl and debtl, capl and gdp, and gdp and turno which are significant at the 5% level. The covariance between the decision and regression random errors is significantly different from zero which implies that governments' decision to privatise is based on some utility maximizing criterion and not a random process. From table 9 it is evident that the random intercepts are significant at the 5% level. Table 11 presents marginal effects of capl, debtl , gdp and turno which are all positive and significant at the 5% level except for turno. These results assert the significant role of market liquidity and debt level in determining privatisation revenue.

Model goodness of fit is gauged using sum of log conditional predictive ordinate (CPO) for cross validation sample. Results presented in table 13 show that the model with more regressors model 1 is preferable to model 2 based on higher sum of log CPO although the difference is not really substantial.

The quantile-quantile (Q-Q) plot in figure 4 shows that the normality assumption of errors is satisfied. The residual plot does not deviate far from the normal line.

Soil Conservation Technology Adoption Study

To demonstrate the flexibility of the MCMC method, that is, it is easily implemented to suit simpler models with least modifications, it is used to analyze agricultural data on adoption of soil conservation technology using a two-regime switching regression model with known threshold value and only fixed effects. Data for this empirical estimation is from a sample of 150 parcels with and without contour hedgerows from a primary survey of 70 farmers from Cebu in central Visayas and 60 farmers from Claveria, Misamis Oriental in the Philippines for the 1995 crop year. Data consists of observations from one parcel per farmer for different seasons and corn variety whenever applicable. The survey sites are upland areas where corn is the main crop.

The decision process of whether or not to adopt a soil conservation technology is specified by the latent adoption decision variable

$$d_j^* = \sum_{k=0}^{13} \gamma_k z_{jk} + \eta_j, \quad j=1,\dots,150 \tag{25}$$

and the farmer's corn yield is specified by the regression model which in this case is the production function considered to be of a general flexible translog form which allows for interaction among variables,

$$y_{j1} = \beta_{01} + \sum_{k=1}^{2} \beta_{k1} V_{jk1} + \sum_{l=1}^{4} \omega_{l1} W_{jl1} + \sum_{k=1}^{2}\sum_{k'=1}^{2} \beta_{kk'1} V_{jk1} V_{jk'1} + \sum_{k=1}^{2}\sum_{l=1}^{4} \upsilon_{kl1} V_{jk1} W_{jl1} + \sum_{p=1}^{4} \varsigma_{p1} D_{p1} + \varepsilon_{j1}$$

for adopters or if $d_j^* > 0$ (26)

$$y_{j0} = \beta_{00} + \sum_{k=1}^{2} \beta_{k0} V_{jk0} + \sum_{l=1}^{4} \omega_{l0} W_{jl0} + \sum_{k=1}^{2}\sum_{k'=1}^{2} \beta_{kk'0} V_{jk0} V_{jk'0} + \sum_{k=1}^{2}\sum_{l=1}^{4} \upsilon_{kl0} V_{jk0} W_{jl0} + \sum_{p=1}^{4} \varsigma_{p0} D_{p0} + \varepsilon_{j0}$$

for non-adopters or if $d_j^* \leq 0$. (27)

V refer to inputs, D are dummy variables and W are condition factors. The input variables for this study are the natural logarithm of labor mandays and the natural logarithm of the total amount of fertilizer applications. The dummy variables are the indicator variable for some form of agricultural training or agricultural extension program participation, location (for either Cebu or Claveria) , corn variety (native and modern hybrid) and dry or wet (rainy) season. The other variables are the farmer characteristics age and education, and farm characteristics soiltype and slope. Table 14 contains the estimates of the adoption decision model and the parameter estimates of regime coefficients with associated standard deviation and 2.5% and 97.5% quantiles, regime error variances and covariances of errors from the adoption decision model and production regime models obtained using reduced MCMC method. The significant determinants of adoption are location, age, education, training, land tenure status, parcel size, soiltype, interaction of slope and location, distance of farm to road, and occurrence of erosion. These results agree with those of Schultz (1975) and Fuglie and Bosch (1995) that some form of formal education or training enable farmers to adjust to technological innovation. Farmers from Cebu and Claveria apparently differ in terms of their adoption decision process. Farmers' productivity depend on whether or not they are adopters or non-adopters. For adopters, the significant explanatory variables for productivity are location, farmer's average household age, education, training, ln of labor in mandays, soiltype, ln of amount of fertilizer applications, season and corn variety. Corn yield increases with the amount of labor and amount of fertilizer input in a quadratic manner. Corn yield also significantly varies with interactions of ln labor and age, ln labor and education, ln labor and soiltype, ln fertilizer and education, ln fertilizer and soiltype, and ln fertilizer and slope. For farmers who choose not to adopt soil conservation technology, the factors significantly influencing productivity are all regressors except the square of fertilizer applications, interactions ln fertilizer and education , and ln fertilizer and slope. Results show that the modern hybrid of corn variety yields more output than the traditional variety for both adopters and non-adopters. From table 15 it is evident that all the coefficients are significantly different for the production regimes of adopters and non-adopters except soiltype, square of ln fertilizer and interactions ln labor and ln fertilizer, ln fertilizer and education. These results indicate that the average productivity effect is significant since the first term in equation (17) is significant. The covariance of the errors from the adoption decision model and the production regime errors are positive for both adopters and non-adopters. Both adopters and non-adopters base their decision on a latent utility maximizing criterion. However, only the covariance for non-adopters is significantly different from zero at the 5% level although the covariance for adopters are almost significant. The same signs of $\sigma_{\eta\varepsilon_1}$ and $\sigma_{\eta\varepsilon_0}$ indicate hierarchical sorting. This implies that farmers who adopt soil conservation technology have above average productivity whether they choose to be an adopter or a non-adopter. But they are more productive if they opt to be adopters rather than non-adopters. On the other hand, farmers who choose not to adopt have below average productivity whether they adopt soil conservation technology or not. However, their productivity is higher from non-adoption. This is not surprising since farmers who adopt soil conservation technologies are better educated and had formal training on soil conservation technology and agricultural extension and they are relatively younger compared to non-adopters. Non-adopters are less educated compared to adopters and only 40% of them had some training on soil conservation technology and/or agricultural extension programs. Table 16 shows that the marginal effect of ln fertilizer of corn yield is higher for adopters of contour hedgerows than for non-adopters. Farmers who adopt soil conservation technology are probably knowledgeable of the optimum input mixture. The productivity differential coefficient or the difference between the error covariances for the production regimes of adopters and non-adopters is 0.2964 and is not significantly different from zero at the 5% level although almost significant. However, it is still economically significant since it implies that the benefit from adoption due to latent productivity attributes is 30%. Thus, the expected gross productivity from adoption which is the sum of the average productivity and differential effect is at least 30%.

The quantile-quantile (Q-Q) plot in figure 5 shows that the normality assumption of errors is satisfied. The residual plot does not deviate far from the normal line.

7. Summary and Conclusion

Estimation and inference procedures for the parameters of the endogenous switching regression model and sample selection models are developed. Random coefficients are incorporated in both the decision and regime regression models to account for heterogeneity across individual units or clusters and correlation within clusters. Tobit type regime regression equations are also considered. A combination of Markov chain Monte Carlo methods, data augmentation and Gibbs sampling is used to facilitate computation of Bayes posterior statistics. From the simulations, we can conclude that for both the full and reduced blocking algorithms, estimation and inference results are fairly robust to prior information and parameter estimates are quite accurate especially for the fixed parameters. However, the MCMC reduced blocking algorithm performs better than the full blocking algorithm both in the simulation study and real data applications in that it rarely gets trapped in absorbing states and it provides models with better fit. It is evident from MCMC methods application to real data sets that the normality assumption still holds for most commonly encountered economic data.

The endogenous switching regression model specification considered in this paper may be extended to include nested or multilevel sources of heterogeneity and more than two regimes. Moreover, error components in the decision and regime equations may be modelled as GARCH errors.

References

Barachina, R., and Engracia, M. (1999) “A New Estimator for Panel Data Sample Selection Models,” *Annales-d’Economie-et-de-Statistique*, 0(55-56), 153-181.

Bortolotti, B., Fantini, M., and Siniscalco, D. (2001) “Privatisation Around the World: New Evidence from Panel Data,” *FEEM Note di Lavoro*, 77:01.

Charlier, E., Melenberg, B, and van Soest, A. (2001) “An Analysis of Housing Expenditure Using Semiparametric Models and Panel Data,” *Journal of Econometrics*, 71:107.

Devroye, L, (1986) *Non Uniform Random Variate Generation*, Springer-Verlag, New York Inc.: New York.

Fuglie, K. and Bosch, D.(1995) “Economic and Environmental Implications of Soil Nitrogen Testing: A Switching Regression Analysis,” *American Journal of Agricultural Economics*, 77, 891-900.

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, S., and Geman, D. J. (1984) “Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration of Images,” *IEEE Transactions on Pattern Analysis Economy*, 82,1119-1143.

Geweke, J. (1991) “Efficient Estimation From Multivariate Normal and Student-t Distributions Subject to Linear Constraints,” *Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface*, 571-578.

Gowrisankaran, G. and Town, R. (1999) "Estimating the Quality of Care in Hospitals Using Instrumental Variables," *Journal of Health Economics*, 18:747-67.

Gronau, R. (1974) “Wage Comparisons: A Selectivity Bias,” and Lewis, H., “Comments on Selectivity Biases in Wage Comparisons,” *Journal of Political Economy*, 82, 1119-1155.

Hartley, M. J. (1978) Comment on “Estimating Mixtures of Normal Distributions and Switching Regression Models,” *Journal of the American Statistical Association*, 73, 738-741.

Hosmer, D. (1973) “On MLE of the Parameters of a Mixture of Two Normal Distributions When the Sample Size is Small” *Communications in Statistics*, 1, 217-227.

Kyriazidou, E. (1997) “Estimation of a Panel Data Sample Selection Model,” *Econometrica* , 65,1335-1364.

Liu, J. S., Wong, W. H., and Kong, A. (1994) “Covariance Structure of the Gibbs Sampler with Applications to the Comparisons of Estimators and Augmentation Schemes,” *Biometrika*, 81, 27-40.

Maddala, G., and Nelson, F. (1975) “Specification Errors in Limited Dependent Variable Models,” *Working Paper 96*, National Bureau of Economic Research, Cambridge.

Mundaca, G. (2001) “Central Bank Interventions and Exchange Rate Band Regimes,” *Journal of International Money and Finance* , 20: 677-700.

Nawata, K., and McAleer, M. (2001) “Size Characteristics of Tests for Sample Selection Bias: A Monte Carlo Comparison and Empirical Example,” *Econometric Reviews*, 20, 105-112.

Odejar, M. E., and McNulty, M. S. (2001) "Bayesian Analysis of the Stochastic Switching Regression Model Using Markov Chain Monte Carlo Methods," *Computational Economics*, 17:265-284.

Quandt, R., Ramsey, J. and James, B. (1978) "Estimating Mixtures of Normal Distributions and Switching Regressions," *Journal of the American Statistical Association*, 73, 730- 738.

Robert, C. (1994) *The Bayesian Choice*, Springer-Verlag New York Inc.: New York.

Ripley, B. (1987) *Stochastic Simulation*, John Wiley: New York.

Schultz, T.W. (1975) "The Value of the Ability to Deal With Disequilibrium," *Journal of Economic Literature*, 13, 827-846.

Tanner, M., and Wong, W. (1987) "The Calculation of Posterior Distributions by Data Augmentation (With Discussion)," *Journal of the American Statistical Association*, 82, 528-550.

Terza, J. (1998) "Endogenous Treatment Effects," *Journal of Econometrics*, 84(1),129-154.

Tierney, L. (1991) "Exploring Posterior Distributions Using Markov Chains," *Computing Science and Statistics* : 23 Symposium on the Interface (Keramidas, E., ed) A.S. A. : Alexandria, Virginia.

Wei, Shang-Jin (1999) "Currency Hedging and Goods Trade," *European Economic Review*, 43, 1371-1394.

Willis, R., and Rosen, S. (1979) "Education and Selection," *Journal of Political Economy*, 87, S7-S36.

Zeger, S. L., and Karim, M. R. (1991) “Generalized Linear Models with Random Effects;a Gibbs Sampling Approach,” *Journal of the American Statistical Association*, 86, 79-86.

Table 1

Algorithm 1 (and Algorithm 2) Simulation Results For Varying Priors

	γ_0	γ_1	γ_2
θ	1.0000	-0.0040	0.3000
$\bar{\hat{\theta}}$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.9659 (0.9673)	-0.0042 (-0.0046)	0.2921 (0.2978)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.9537 (0.9601)	-0.0027 (-0.0020)	0.2874 (0.2898)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.9297 (0.9461)	-0.0018 (-0.0027)	0.2715 (0.2784)
$s_{\hat{\theta}}$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0526 (0.0517)	0.0077 (0.0080)	0.0410 (0.0516)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.0502 (0.0496)	0.0086 (0.0083)	0.0522 (0.0545)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0440 (0.0414)	0.0089 (0.0083)	0.0423 (0.0530)
$\bar{\hat{s}}_\theta$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0621 (0.0474)	0.0184 (0.0185)	0.0478 (0.0474)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.0582 (0.0456)	0.0184 (0.0185)	0.0463 (0.0459)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0538 (0.0436)	0.0184 (0.0184)	0.0441 (0.0440)
Mean Absolute Deviation from True Parameter			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0506 (0.0488)	0.0061 (0.0064)	0.0313 (0.0412)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.0562 (0.0530)	0.0072 (0.0069)	0.0430 (0.0451)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0730 (0.0584)	0.0074 (0.0069)	0.0422 (0.0459)
Coverage of Nominal 95% Interval			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	100 (100)	100 (100)	100 (100)

$\beta_p = 0.7\beta$, Ну = $0.7\Sigma_{y_i}$, Нк= $0.7\Sigma_\kappa$	100 (100)	100 (100)	100 (100)
$\beta_p = 0.5\beta$, Ну = $0.5\Sigma_{y_i}$, Нк= $0.5\Sigma_\kappa$	100 (100)	100 (100)	100 (100)

Table 2

Algorithm 1 (and Algorithm 2) Simulation Results For Varying Priors

	β_{10}	β_{11}	β_{12}
θ	14.0000	-0.3000	0.5300
$\bar{\hat{\theta}}$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	13.8958 (13.9121)	-0.2836 (-0.3028)	0.4246 (0.4334)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	13.8923 (13.8877)	-0.2872 (-0.2768)	0.4192 (0.4042)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	13.8823 (13.8953)	-0.2652 (-0.2768)	0.4040 (0.3971)
$s_{\hat{\theta}}$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.1083 (0.0977)	0.0756 (-0.0746)	0.0828 (0.0743)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.1065 (0.1060)	0.0768 (0.0862)	0.0677 (0.0726)
$\beta_p = 0.5\ \beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.1170 (0.1060)	0.0807 (0.0865)	0.0781 (0.0751)
$\bar{\hat{s}}_\theta$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.1138 (0.1044)	0.0883 (0.0879)	0.0928 (0.0916)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.1067 (0.0985)	0.0851 (0.0841)	0.0889 (0.0875)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0984 (0.0920)	0.0804 (0.0795)	0.0833 (0.0823)
Mean Absolute Deviation from True Parameter			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.1204 (0.1043)	0.0599 (0.0619)	0.1124 (0.1029)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.1224 (0.1245)	0.0628 (0.0736)	0.1144 (0.1284)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.1385 (0.1226)	0.0663 (0.0731)	0.1307 (0.1352)
Coverage of Nominal 95% Interval			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	100 (100)	100 (100)	100 (100)

$\beta_p = 0.7\beta$, Ну = 0.7 Σ_{y_i}, Нк= 0.7 Σ_κ	100 (100)	100 (100)	100 (100)
$\beta_p = 0.5\beta$, Ну = 0.5 Σ_{y_i}, Нк= 0.5 Σ_κ	100 (100)	100 (100)	100 (100)

Table 3

Algorithm 1 (and Algorithm 2) Simulation Results For Varying Priors

	β_{00}	β_{01}	β_{02}
θ	-1.0000	0.7000	-0.3000
$\bar{\bar{\theta}}$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	-0.9072 (-0.8748)	0.6896 (0.6716)	-0.0669 (-0.0659)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	-0.8423 (-0.8405)	0.6655 (0.6833)	-0.0614 (-0.0624)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	-0.7898 (-0.8035)	0.6580 (0.6508)	-0.0579 (-0.0560)
$s_{\hat{\theta}}$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0866 (0.1004)	0.0922 (0.1032)	0.0194 (0.0181)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.1036 (0.1054)	0.0911 (0.0928)	0.0185 (0.0164)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0928 (0.1021)	0.1048 (0.1002)	0.0220 (0.0192)
$\bar{\hat{s}}_\theta$			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.1147 (0.1045)	0.0972 (0.0957)	0.0569 (0.0569)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.1052 (0.0952)	0.0898 (0.0892)	0.0518 (0.0517)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0921 (0.0846)	0.0790 (0.0785)	0.0463 (0.0461)
Mean Absolute Deviation from True Parameter			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$ $H_\kappa = 0.9\Sigma_\kappa$	0.1063 (0.1363)	0.0744 (0.0872)	0.2331 (0.2341)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.1618 (0.1667)	0.0799 (0.0752)	0.2387 (0.2376)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.2108 (0.2000)	0.0911 (0.0898)	0.2421 (0.2440)
Coverage of Nominal 95% Interval			
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	100 (100)	100 (100)	100 (100)

$\beta_p = 0.7\beta$, Ну = $0.7\Sigma_{y_i}$, Нк= $0.7\Sigma_\kappa$	100 (100)	100 (100)	100 (100)
$\beta_p = 0.5\beta$, Ну = $0.5\Sigma_{y_i}$, Нк= $0.5\Sigma_\kappa$	100 (100)	100 (100)	100 (100)

Table 4

Algorithm 1 (and Algorithm 2) Simulation Results For Varying Priors

	$\sigma^2_{\varepsilon_1}$	$\sigma^2_{\varepsilon_0}$	$\sigma_{\eta\varepsilon_1}$	$\sigma_{\eta\varepsilon_0}$
θ	9.0000	2.0000	0.4000	-0.2000
$\bar{\bar{\theta}}$				
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	8.5771 (8.3950)	1.9207 (1.8901)	0.2324 (0.2197)	-0.1153 (-0.1108)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	7.6522 (7.4778)	1.5724 (1.5481)	0.1857 (0.1736)	-0.0875 (-0.0840)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	6.6169 (6.5100)	1.2150 (1.2117)	0.1364 (0.1264)	-0.0612 (-0.0591)
$s_{\hat{\theta}}$				
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.2205 (0.2330)	0.0381 (0.0378)	0.0131 (0.0133)	0.0044 (0.0055)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.2195 (0.2158)	0.0332 (0.0457)	0.0115 (0.0139)	0.0048 (0.0047)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.2037 (0.1777)	0.0344 (0.0471)	0.0106 (0.0116)	0.0047 (0.0049)
$\bar{\hat{s}}_\theta$				
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.2949 (0.2821)	0.0839 (0.0816)	0.0835 (0.0830)	0.0391 (0.0391)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.2628 (0.2514)	0.0689 (0.0668)	0.0772 (0.0765)	0.0343 (0.0342)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.2272 (0.2185)	0.0531 (0.0524)	0.0702 (0.0699)	(0.1388) (0.0293)
Mean Absolute Deviation from True Parameter				
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.4263 (0.6058)	0.0794 (0.1099)	0.1675 (0.1803)	0.0847 (0.0892)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	1.3477 (1.5222)	0.4276 (0.4519)	0.2143 (0.2263)	0.1125 (0.1160)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	2.3831 (2.4900)	0.7850 (0.7883)	0.2636 (0.2736)	0.1388 (0.1409)
Coverage of Nominal 95% Interval				
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	100 (100)	100 (100)	100 (100)	100 (100)

$\beta_p = 0.7\beta$, Ну = 0.7 Σ_{y_i}, Нк= 0.7 Σ_κ	100 (100)	100 (100)	100 (100)	100 (100)
$\beta_p = 0.5\beta$, Ну = 0.5 Σ_{y_i}, Нк= 0.5 Σ_κ	100 (100)	100 (100)	100 (100)	100 (100)

Table 5

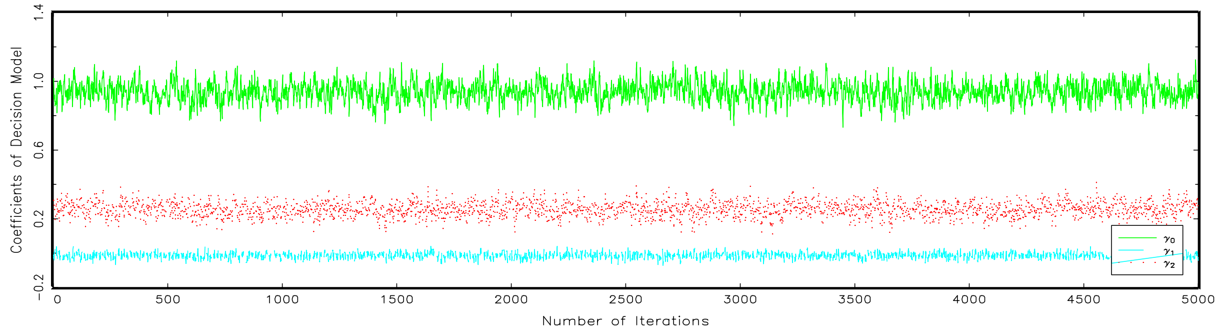
Algorithm 1 (and Algorithm 2) Simulation Results For Varying Priors

	$\sigma^2_{\tau_0}$	$\sigma^2_{\tau_1}$	$\sigma^2_{\kappa_{10}}$	$\sigma^2_{\kappa_{11}}$	$\sigma^2_{\kappa_{00}}$	$\sigma^2_{\kappa_{01}}$
θ	0.1000	0.0700	0.1000	0.0600	0.1000	0.0500
$\bar{\hat{\theta}}$						
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0796 (0.0781)	0.0588 (0.0582)	0.0882 (0.0887)	0.0532 (0.0535)	0.0896 (0.0886)	0.0415 (0.0415)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.0650 (0.0635)	0.0476 (0.0475)	0.0698 (0.0700)	0.0420 (0.0421)	0.0715 (0.0712)	0.0327 (0.0326)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0496 (0.0488)	0.0362 (0.0359)	0.0506 (0.0509)	0.0302 (0.0302)	0.0536 (0.0533)	0.0237 (0.0237)
$s_{\hat{\theta}}$						
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0032 (0.0031)	0.0026 (0.0019)	0.0020 (0.0021)	0.0009 (0.0009)	0.0024 (0.0022)	0.0005 (0.0004)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.0033 (0.0024)	0.0023 (0.0019)	0.0018 (0.0015)	0.0008 (0.0008)	0.0022 (0.0025)	0.0004 (0.0005)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0027 (0.0022)	0.0019 (0.0016)	0.0011 (0.0013)	0.0005 (0.0005)	0.0025 (0.0028)	0.0004 (0.0004)
$\bar{\hat{s}}_\theta$						
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0141 (0.0136)	0.0109 (0.0107)	0.0174 (0.0176)	0.0106 (0.0107)	0.0179 (0.0175)	0.0077 (0.0077)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.0118 (0.0114)	0.0091 (0.0090)	0.0140 (0.0141)	0.0085 (0.0085)	0.0146 (0.0144)	0.0062 (0.0061)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0095 (0.0092)	0.0073 (0.0071)	0.0104 (0.0104)	0.0062 (0.0062)	0.0114 (0.0113)	0.0045 (0.0045)
Mean Absolute Deviation from True Parameter						
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	0.0204 (0.0218)	0.0112 (0.0118)	0.0117 (0.0113)	0.0068 (0.0065)	0.0104 (0.0114)	0.0084 (0.0085)
$\beta_p = 0.7\beta$, $H_y = 0.7\Sigma_{y_i}$, $H_\kappa = 0.7\Sigma_\kappa$	0.0350 (0.0365)	0.0223 (0.0225)	0.0302 (0.0299)	0.0180 (0.0179)	0.0285 (0.0288)	0.0173 (0.0174)
$\beta_p = 0.5\beta$, $H_y = 0.5\Sigma_{y_i}$, $H_\kappa = 0.5\Sigma_\kappa$	0.0504 (0.0511)	0.0338 (0.0341)	0.0494 (0.0491)	0.0298 (0.0298)	0.0464 (0.0467)	0.0262 (0.0263)
Coverage of Nominal 95% Interval						
$\beta_p = 0.9\beta$, $H_y = 0.9\Sigma_{y_i}$, $H_\kappa = 0.9\Sigma_\kappa$	100	100	100	100	100	100

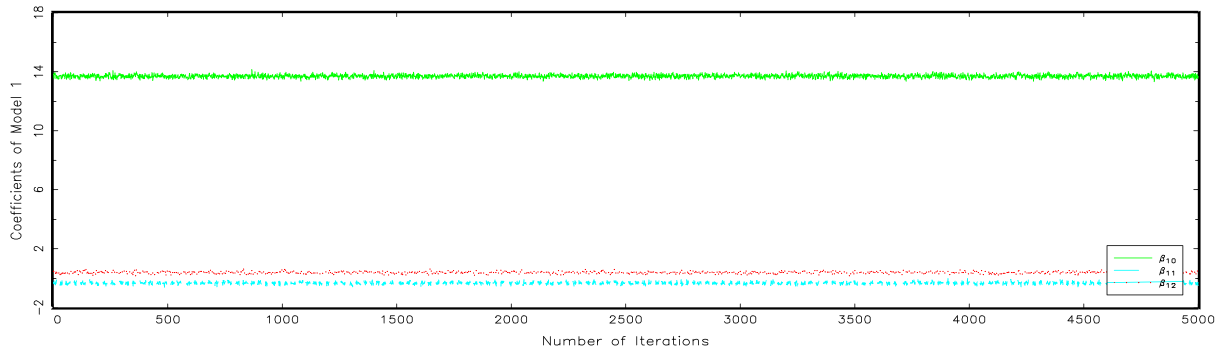
	(100)	(100)	(100)	(100)	(100)	(100)
$\beta_p = 0.7\beta$, Ну = 0.7Σ_{y_i}, Нк= 0.7Σ_κ	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)
$\beta_p = 0.5\beta$, Ну = 0.5Σ_{y_i}, Нк= 0.5Σ_κ	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)	100 (100)

Figure 1a

Iteration Results for Decision Model



Iteration Results for Model 1



Iteration Results for Model 0

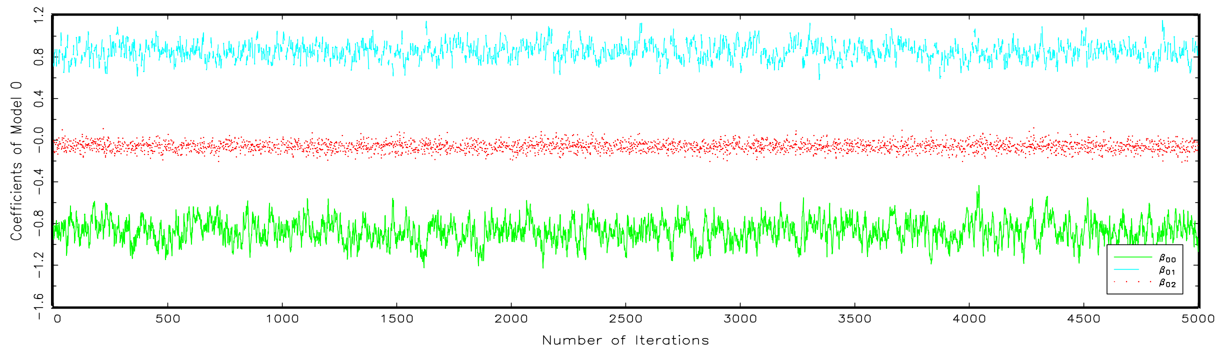
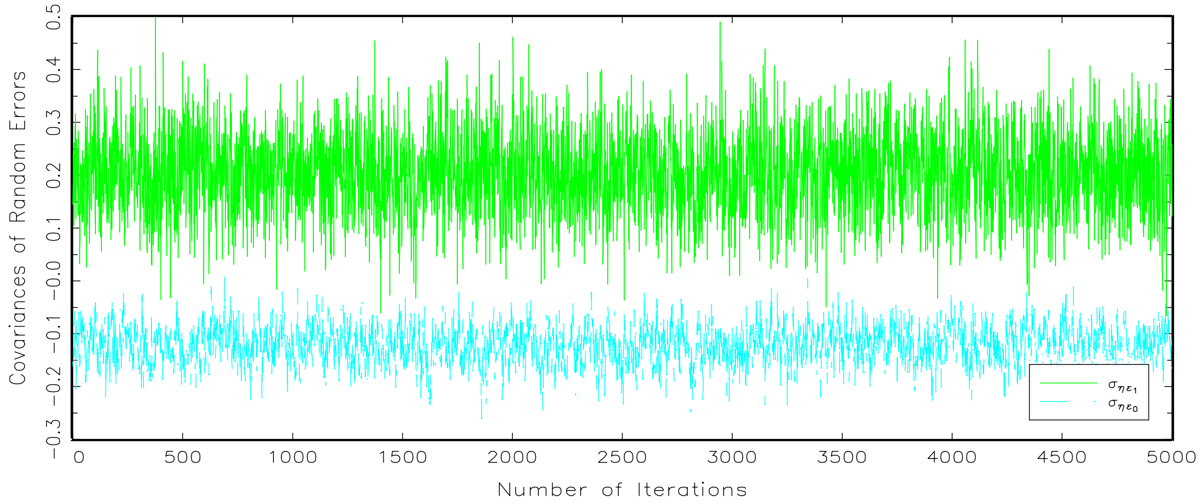


Figure 1b

Iteration Results for Covariances of Random Errors



Iteration Results for Variances of Random Errors

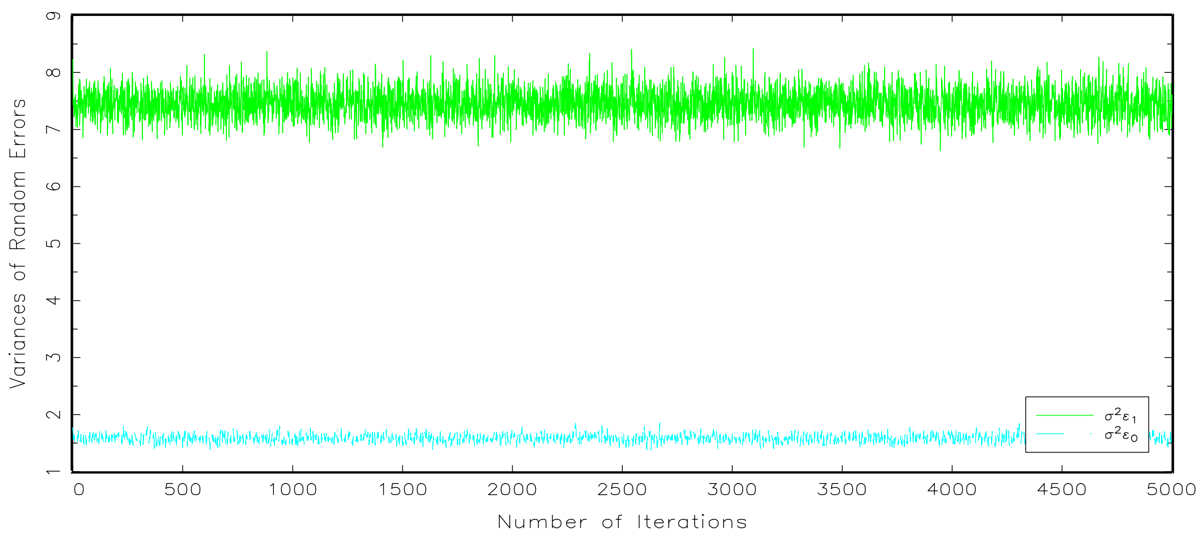
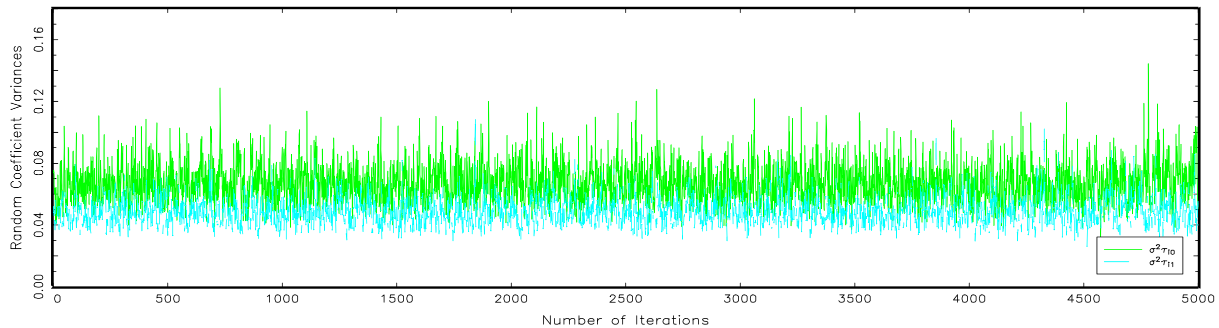
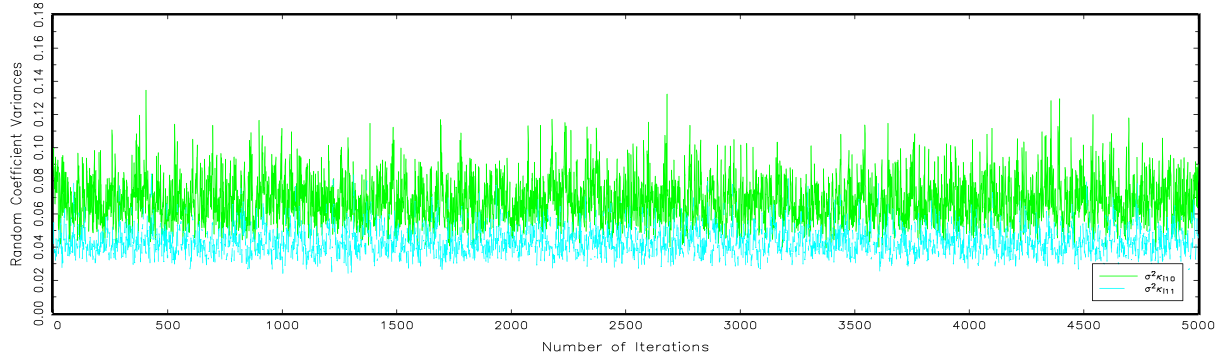


Figure 1c

Iteration Results for Variances of Random Coefficients for the Decision Model



Iteration Results for Variances of Random Coefficients for Model 1



Iteration Results for Variances of Random Coefficients for Model 0

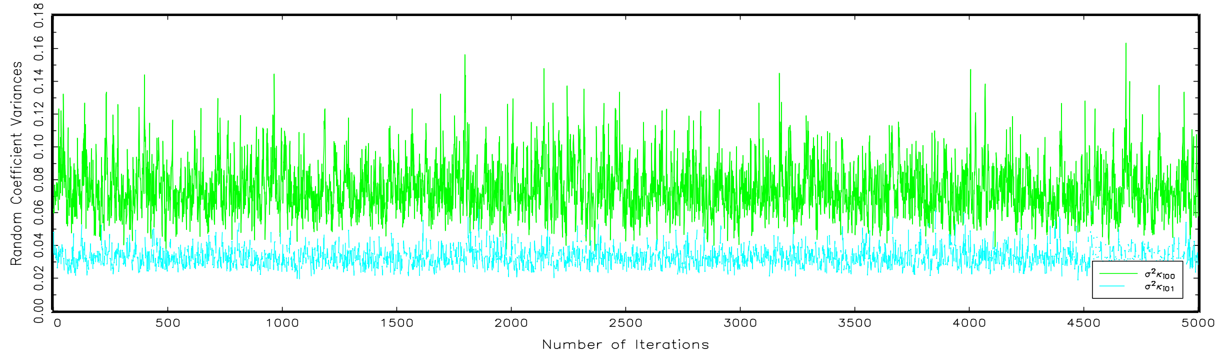
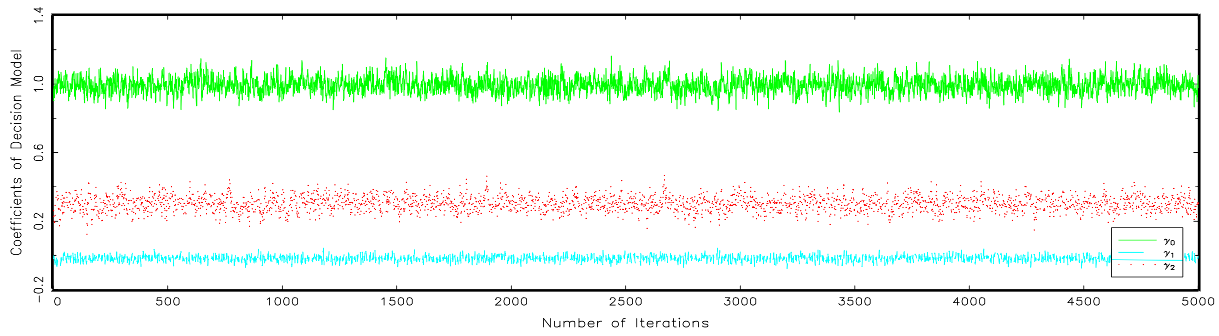
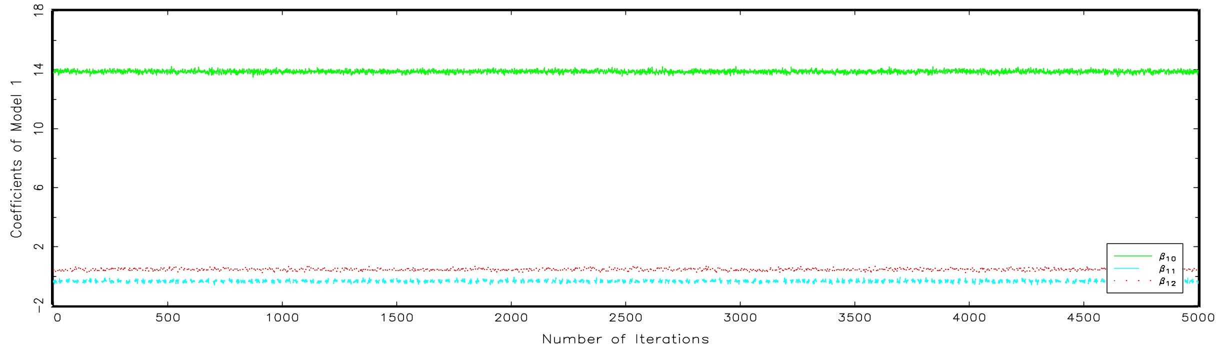


Figure 2a

Iteration Results for Decision Model



Iteration Results for Model 1



Iteration Results for Model 0

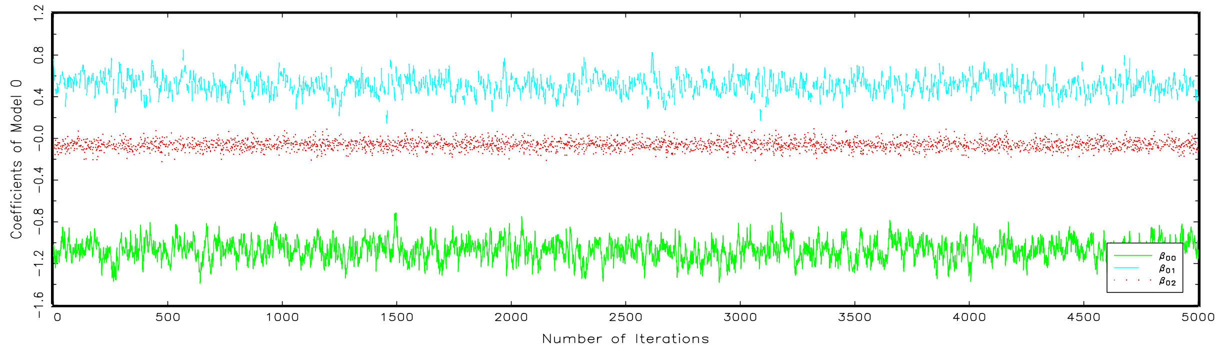
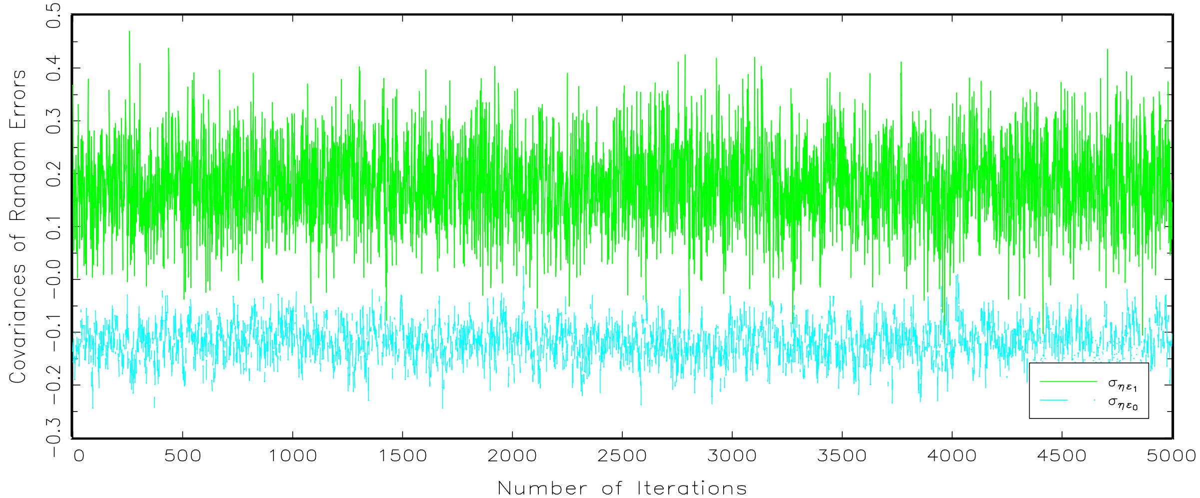


Figure 2b

Iteration Results for Covariances of Random Errors



Iteration Results for Variances of Random Errors

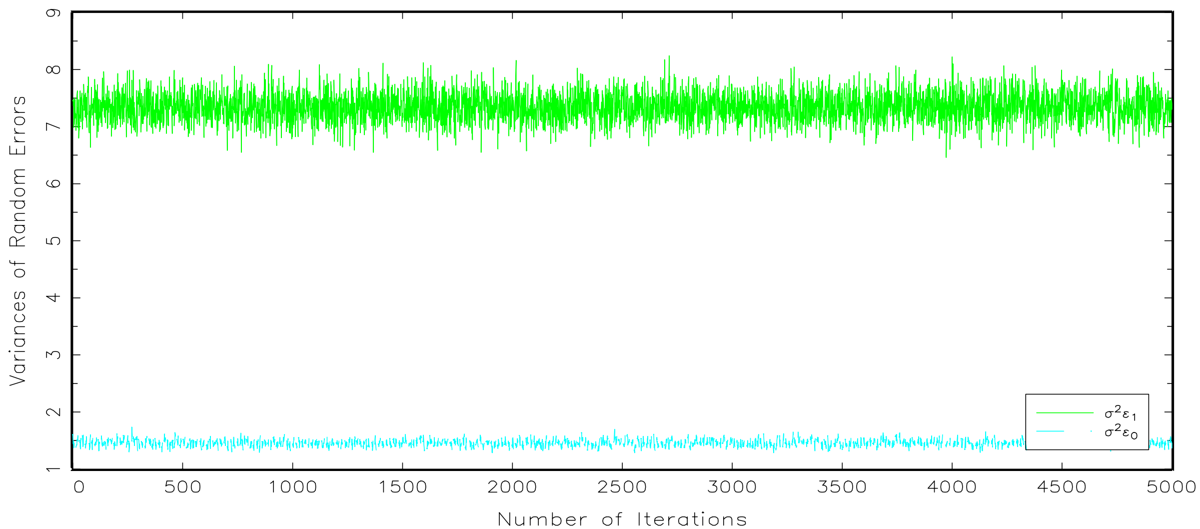
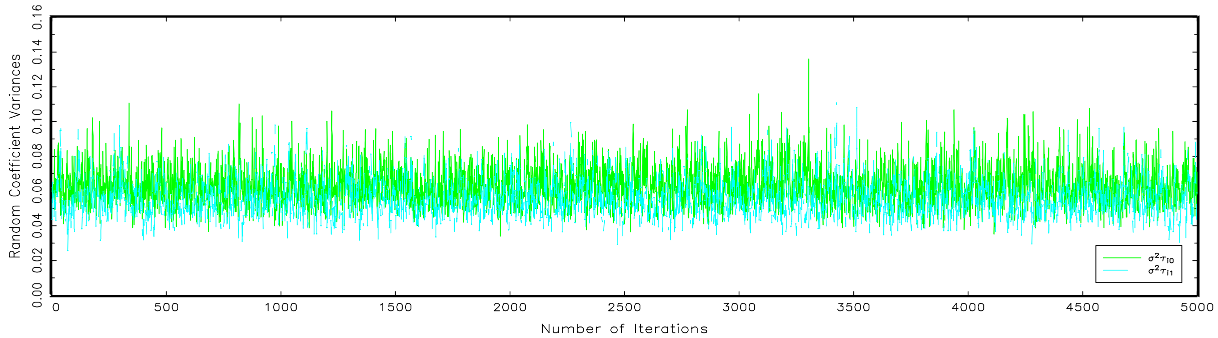
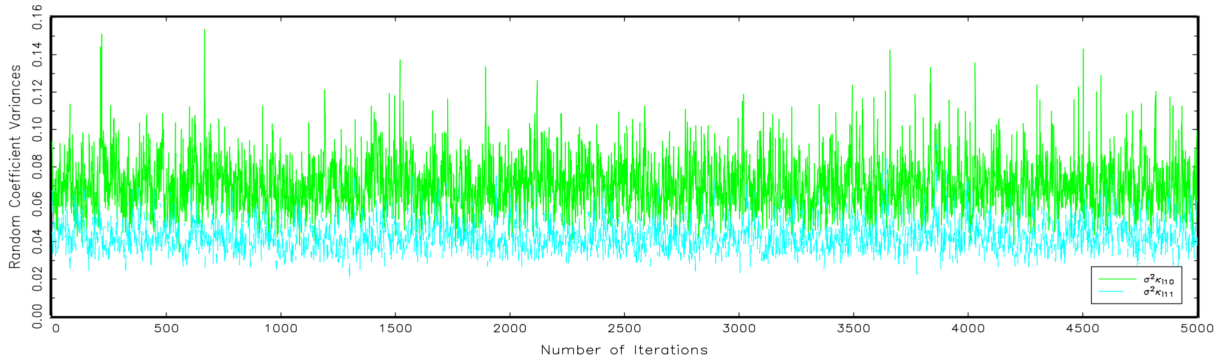


Figure 2c

Iteration Results for Variances of Random Coefficients for the Decision Model



Iteration Results for Variances of Random Coefficients for Model 1



Iteration Results for Variances of Random Coefficients for Model 0

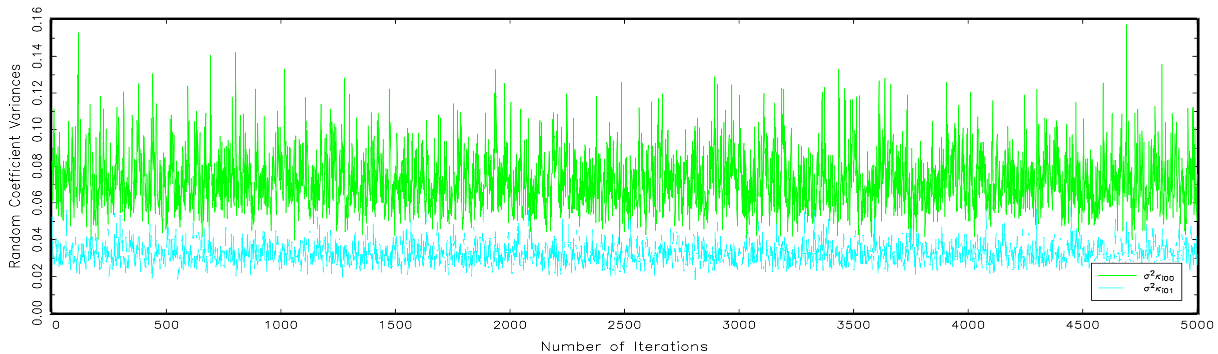


Figure 3. Quantile-Quantile Plot of Residuals for Currency Hedging and Goods Trade Data

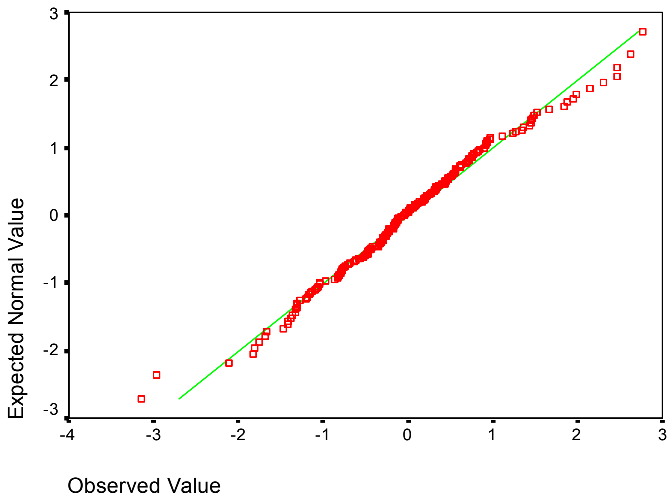


Table 6

Estimates for the Currency Hedging and Goods Trade Study

Decision Model Variables	**Coefficients**	**Standard Error**	**2.5% Quantile**	**97.5% Quantile**
dist	-0.6644	0.0231	-0.7089	-0.6176
gnpc	0.1413	0.0120	0.1180	0.1652
comlang	0.5773	0.0518	0.4770	0.6801
xrvol	3.6863	0.0803	3.5308	3.8454
Regression Model Variables for Non-Hedgers	**Coefficients**	**Standard Error**	**2.5% Quantile**	**97.5% Quantile**
dist	-0.4952	0.0187	-0.5314	-0.4590
gnp	0.8028	0.0083	0.7865	0.8190
comlang	0.6180	0.0501	0.5208	0.7169
xrvol	2.2861	0.0549	2.1782	2.3934
Regression Model Variables for Hedgers	**Coefficients**	**Standard Error**	**2.5% Quantile**	**97.5% Quantile**
dist	-0.5809	0.0215	-0.6227	-0.5381
gnp	0.7302	0.0116	0.7079	0.7535
comlang	0.5977	0.0508	0.4997	0.6966
xrvol	-28.7620	0.3201	-29.4302	-28.1562
Error Variance-Covariance	**Posterior Mean**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
$\sigma_{\eta\varepsilon_0}$	0.7096	0.3225	0.0744	1.3442
$\sigma_{\eta\varepsilon_1}$	-0.4141	0.2032	-0.8299	-0.0297
$\sigma^2_{\varepsilon_0}$	11.0916	1.1415	9.0450	13.5353
$\sigma^2_{\varepsilon_1}$	3.6103	0.5232	2.6637	4.7319
$\sigma^2_{\tau_0}$	0.3609	0.0743	0.2405	0.5290
$\sigma^2_{\kappa_{00}}$	0.3772	0.0847	0.2485	0.5732
$\sigma^2_{\kappa_{10}}$	0.9546	0.1913	0.6407	1.3956

Table 7

Marginal Effects Estimates for the Currency Hedging and Goods Trade Study

Variables for Non-Hedgers	**Posterior Mean**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
Dist	-0.3066	0.1376	-0.5314	-0.4590
Gnp	-28.7620	0.3201	2.1782	2.3934
Xrvol	0.3593	0.1290	0.5208	0.7169
Variables for Hedgers	**Posterior Mean**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
Dist	-0.4952	0.0187	-0.5674	-0.0246
Gnp	2.2861	0.0549	-29.4302	-28.1562
Xrvol	0.6180	0.0501	0.0946	0.5994

Table 8

Estimates of Difference for Hedgers and Non-Hedgers for the Currency Hedging and Goods Trade Study

Regression Model Variables	**Coefficient Difference (Posterior Mean)**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
dist	-0.0857	0.0284	-0.1406	-0.0290
gnp	-0.0727	0.0143	-0.1002	-0.0444
comlang	-0.0203	0.0716	-0.1621	0.1178
xrvol	-31.0481	0.3256	-31.7271	-30.4290
Error Variance-Covariance	**Posterior Mean**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
$\sigma_{\eta\varepsilon_0} - \sigma_{\eta\varepsilon_1}$	-1.1237	0.3864	-1.8911	-0.3672

Figure 4. Quantile-Quantile Plot of Residuals for Cross-Country Privatisation Data

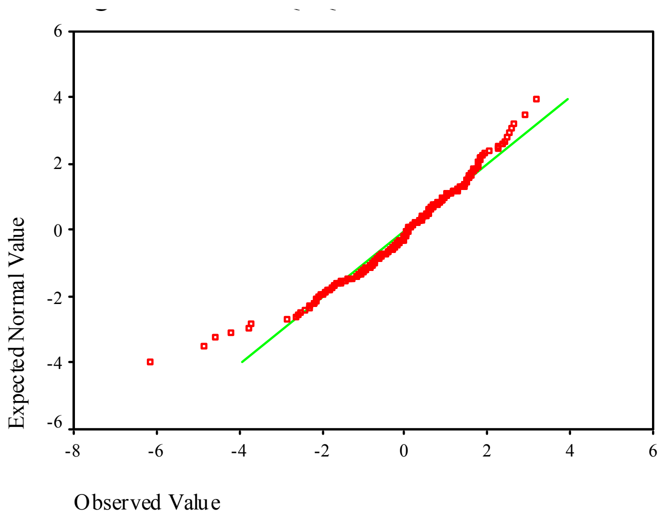


Table 9

Estimates for the Cross-Country Privatisation Study Model 1

Decision Model Variables	**Coefficients**	**Standard Error**	**2.5% Quantile**	**97.5% Quantile**
capl	0.3902	0.0733	0.2448	0.5321
debtl	0.0369	0.0155	0.0064	0.0680
gdp	0.0471	0.0134	0.0219	0.0742
turno	0.4160	0.0961	0.2337	0.6033
Nondem	-0.2483	0.2495	-0.7449	0.2375
Right	0.4641	0.0937	0.2830	0.6495
Regression Model Variables	**Coefficients**	**Standard Error**	**2.5% Quantile**	**97.5% Quantile**
capl	1.8524	0.6053	0.6580	3.0590
debtl	0.4692	0.1669	0.1483	0.7950
gdp	0.4915	0.1914	0.1077	0.8707
turno	-1.6937	0.7808	-3.1476	-0.1593
$(\text{capl})^2$	-0.1158	0.0343	-0.1848	-0.0502
$(\text{debtl})^2$	-0.0608	0.0129	-0.0863	-0.0360
$(\text{gdp})^2$	0.0147	0.0140	-0.0130	0.0431
$(\text{turno})^2$	-0.1580	0.0560	-0.2672	-0.0472
(capl) x (debtl)	-0.0734	0.0166	-0.1055	-0.0408
(capl) x (gdp)	-0.0537	0.0493	-0.1520	0.0424
(capl) x (turno)	0.0436	0.0543	-0.0624	0.1480
(debtl) x (gdp)	0.0326	0.0111	0.0113	0.0541
(debtl) x (turno)	0.0136	0.0088	-0.0037	0.0310
(gdp) x (turno)	0.1530	0.0645	0.0246	0.2750
$\sigma_{\eta\varepsilon_1}$	-1.1486	0.1418	-1.4321	-0.8810
$\sigma^2_{\varepsilon_1}$	4.5939	0.3466	3.9778	5.3039
$\sigma^2_{\tau_{i0}}$	2.7188	0.6064	1.7742	4.1133
$\sigma^2_{\kappa_{i10}}$	10.1260	2.2420	6.6528	15.2328

Table 10

Estimates for the Cross-Country Privatisation Study Model 2

Decision Model Variables	**Coefficients**	**Standard Error**	**2.5% Quantile**	**97.5% Quantile**
debtl	0.0096	0.0153	-0.0209	0.0382
gdp	-0.0088	0.0101	-0.0283	0.0113
Right	0.5451	0.1248	0.3075	0.7959
Regression Model Variables	**Coefficients**	**Standard Error**	**2.5% Quantile**	**97.5% Quantile**
gdp	0.5856	0.0333	0.5249	0.6556
turno	-1.2918	0.8869	-3.0739	0.3045
(capl) x (debtl)	-0.0559	0.0142	-0.0836	-0.0282
(capl) x (turno)	-0.0620	0.0791	-0.2222	0.0948
(gdp) x (turno)	0.1182	0.0752	-0.0215	0.2681
$\sigma_{\eta\varepsilon_1}$	-1.5135	0.1873	-1.9283	-1.1870
$\sigma^2{}_{\varepsilon_1}$	4.8908	0.6271	3.9415	6.3575
$\sigma^2{}_{\tau_0}$	3.0086	0.6882	1.9141	4.6425
$\sigma^2{}_{\kappa_{10}}$	9.2238	2.0087	6.1601	14.0787

Table 11

Marginal Effects Estimates for the Cross-Country Privatisation Study Model 1

Variables	Posterior Mean	Standard Deviation	2.5% Quantile	97.5% Quantile
Capl	1.1072	0.1793	0.7569	1.4633
Debtl	0.8681	0.1284	0.6186	1.1185
Gdp	0.7725	0.1227	0.5271	1.0148
Turno	0.1444	0.2331	-0.3287	0.5967

Table 12

Marginal Effects Estimates for the Cross-Country Privatisation Study Model 2

Variables	Posterior Mean	Standard Deviation	2.5% Quantile	97.5% Quantile
capl	-0.1837	0.0613	-0.3044	-0.0650
debtl	0.1010	0.0308	0.0453	0.1661
gdp	0.3498	0.0723	0.2115	0.4954
turno	-0.1156	0.4993	-1.1662	0.7908

Table 13

Model Comparison for the Cross-Country Privatisation Study

Model	**Sum of Log Conditional Predictive Ordinate for Cross-Validation Sample**	
	Algorithm 1	**Algorithm 2**
Model 1	-86.4425	-74.9771
Model 2	-94.1465	-77.1314

Figure 5. Quantile-Quantile Plot of Residuals for Adoption of Soil-Conservation Technology Data

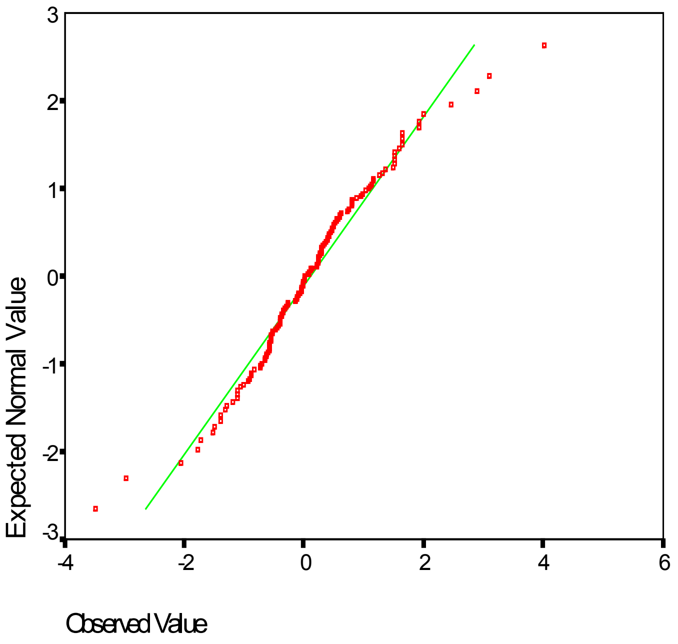


Table 14

Estimates for the Soil Conservation Technology Adoption Study

Decision Model Variables	**Coefficients (Posterior Mean)**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
Constant	1.4928	0.0332	1.4286	1.5585
Location	-1.9055	0.0335	-1.9715	-1.8396
Age (years)	-0.0860	0.0081	-0.1024	-0.0706
Education (years)	0.0676	0.0254	0.0180	0.1185
Ln (Labor)	0.0616	0.0329	-0.0030	0.1251
Training	2.0037	0.0330	1.9385	2.0693
Tenure status	0.2021	0.0332	0.1381	0.2668
Ln(Parcel Size)	0.3836	0.0325	0.3199	0.4469
Soiltype	-0.1893	0.0325	-0.2544	-0.1251
Slope (percent)	0.0081	0.0106	-0.0121	0.0294
Distance to Road	0.4657	0.0330	0.4008	0.5308
Erosion	0.9699	0.0330	0.9056	1.0352
Location x Slope	0.0695	0.0112	0.0486	0.0916
Regression Model Variables for Adopters	**Coefficients (Posterior Mean)**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
Constant	1.9875	0.0302	1.9292	2.0465
Location	0.5945	0.0302	0.5361	0.6530
Age	0.0539	0.0186	0.0175	0.0915
Education	0.1153	0.0284	0.0597	0.1715
Training	0.3845	0.0301	0.3260	0.4439
Ln(Labor)	0.8681	0.0298	0.8098	0.9250
Slope	0.0204	0.0233	-0.0251	0.0672
Soiltype	-0.5348	0.0301	-0.5939	-0.4776
Ln(Fertilizer)	0.1529	0.0302	0.0941	0.2116
Season	-0.2181	0.0303	-0.2766	-0.1599
Variety	0.6871	0.0300	0.6277	0.7451
Ln^2 (Labor)	0.1261	0.0276	0.0709	0.1810
Ln^2 (Fertilizer)	0.0309	0.0094	0.0123	0.0498
Ln(Labor) x Ln(Fertilizer)	0.0314	0.0204	-0.0070	0.0717
Ln(Labor) x Age	-0.0363	0.0063	-0.0493	-0.0241
Ln(Labor) x Education	-0.0391	0.0127	-0.0640	-0.0139

Ln(labor) x Soiltype	0.0859	0.0274	0.0332	0.1410
Ln(Labor) x Slope	0.0004	0.0072	-0.0139	0.0146
Ln(fertilizer) x Age	0.0015	0.0018	-0.0021	0.0052
Ln(Fertilizer) x Education	-0.0150	0.0072	-0.0294	-0.0009
Ln(Fertilizer) x Soiltype	0.0679	0.0243	0.0207	0.1157
Ln(Fertilizer) x Slope	-0.0060	0.0030	-0.0120	-0.0002
Regression Model Variables for Non-Adopters	**Coefficients (Posterior Mean)**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
Constant	8.9927	0.0298	8.9345	9.0526
Location	0.2914	0.0296	0.2330	0.3506
Age	-0.1869	0.0187	-0.2234	-0.1509
Education	-0.2274	0.0291	-0.2861	-0.1716
Training	-0.3940	0.0302	-0.4528	-0.3353
Ln(Labor)	-1.0341	0.0301	-1.0920	-0.9753
Slope	0.2184	0.0244	0.1698	0.2667
Soiltype	-0.5284	0.0294	-0.5866	-0.4704
Ln(Fertilizer)	0.3778	0.0293	0.3207	0.4352
Season	0.1850	0.0299	0.1257	0.2419
Variety	0.9838	0.0298	0.9255	1.0424
Ln^2 (Labor)	0.2297	0.0277	0.1762	0.2860
Ln^2 (Fertilizer)	0.0103	0.0095	-0.0079	0.0289
Ln(Labor) x Ln(Fertilizer)	0.0644	0.0182	0.0295	0.1026
Ln(Labor) x Age	0.0289	0.0055	0.0180	0.0395
Ln(Labor) x Education	0.0342	0.0132	0.0088	0.0606
Ln(labor) x Soiltype	0.5635	0.0278	0.5081	0.6190
Ln(Labor) x Slope	-0.0754	0.0067	-0.0888	-0.0620
Ln(fertilizer) x Age	-0.0049	0.0014	-0.0079	-0.0022
Ln(Fertilizer) x Education	-0.0058	0.0067	-0.0191	0.0073
Ln(Fertilizer) x Soiltype	-0.3467	0.0254	-0.3964	-0.2946
Ln(Fertilizer) x Slope	0.0022	0.0019	-0.0015	0.0062
Error Variance-Covariance	**Posterior Mean**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
$\sigma_{\eta\varepsilon_1}$	0.1712	0.1671	-0.1385	0.5297
$\sigma_{\eta\varepsilon_0}$	0.4676	0.1704	0.1529	0.8349
$\sigma^2_{\varepsilon_1}$	1.3527	0.2087	1.0014	1.8334
$\sigma^2_{\varepsilon_0}$	1.2267	0.1791	0.9234	1.6404

Table 15

Difference of Estimates for Adopters and Non-Adopters of Soil Conservation Technology

Regression Model Variables	**Coefficient Difference (Posterior Mean)**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
Constant	-7.0052	0.0424	-7.0885	-6.9205
Location	0.3032	0.0420	0.2202	0.3860
Age	0.2408	0.0266	0.1887	0.2939
Education	0.3427	0.0411	0.2629	0.4249
Training	0.7785	0.0425	0.6957	0.8627
Ln(Labor)	1.9022	0.0420	1.8183	1.9842
Slope	-0.1980	0.0335	-0.2641	-0.1297
Soiltype	-0.0064	0.0418	-0.0891	0.0744
Ln(Fertilizer)	-0.2249	0.0421	-0.3072	-0.1433
Season	-0.4031	0.0429	-0.4853	-0.3181
Variety	-0.2968	0.0427	-0.3809	-0.2141
Ln^2 (Labor)	-0.1036	0.0384	-0.1833	-0.0275
Ln^2 (Fertilizer)	0.0205	0.0142	-0.0074	0.0484
Ln(Labor) x Ln(Fertilizer)	-0.0329	0.0273	-0.0859	0.0202
Ln(Labor) x Age	-0.0652	0.0084	-0.0819	-0.0485
Ln(Labor) x Education	-0.0733	0.0183	-0.1091	-0.0374
Ln(labor) x Soiltype	-0.4775	0.0384	-0.5523	-0.4007
Ln(Labor) x Slope	0.0758	0.0099	0.0564	0.0953
Ln(fertilizer) x Age	0.0064	0.0024	0.0019	0.0111
Ln(Fertilizer) x Education	-0.0093	0.0101)	-0.0291	0.0109
Ln(Fertilizer) x Soiltype	0.4146	0.0343	0.3448	0.4817
Ln(Fertilizer) x Slope	-0.0082	0.0036	-0.0158	-0.0014
Error Covariance	**Posterior Mean**	**Standard Deviation**	**2.5% Quantile**	**97.5% Quantile**
$\sigma_{\eta\varepsilon_1} - \sigma_{\eta\varepsilon_0}$	-0.2964	0.1682	-0.6351	0.0332

Table 16

Marginal Effects for the Soil Conservation Technology Adoption Study

Variables	Posterior Mean	Standard Deviation	2.5% Quantile	97.5% Quantile
For Adopters				
Ln(Labor)	0.1967	0.1839	-0.1807	0.5561
Ln(Fertilizer)	0.3483	0.0857	0.1817	0.5223
For Non-Adopters				
Ln(Labor)	1.2366	0.2028	0.8498	1.6313
Ln(Fertilizer)	0.0045	0.0885	-0.1648	0.1780