Bayesian Panel Data Model with Constraints

Ameera Jaber Mohaisen and Saja Yaseen Abdulsamad
Mathematics Department College of Education for Pure Science AL-Basrah University-Iraq

E-mail:ameera.jaber@yahoo.com
E-mail:saja.shareeda@yahoo.com

Abstract. In this paper, we consider a set of linear constraints on the coefficients of the random panel data model. Furthermore, Bayesian approach based on Markov Chain Monte Carlo (MCMC) is employed to making inferences on the model coefficients subject to this constraints.

Keywords. Panel Data Model, Maximum Likelihood Method, constrained maximum likelihood estimator, Prior distribution, Posterior distribution, Bayesian estimation.

1. Introduction

The analysis of panel data allows the model builder to learn about economic processes while accounting for both heterogeneity across individuals, firms, countries, and so on and for dynamic effects that are not visible in cross sections. Modelling in this context often calls for complex stochastic specifications.[7]. The panel data model has been investigated by the many researcher as Elhorst in (2001) presented paper surveys panel data models extended to spatial error autocorrelation or spatially lagged dependent variable, [5]. Hurlin in (2004) proposed a simple test of Granger (1969) non causality hypothesis in heterogeneous panel data models with fixed coefficients,[10]. Bun , a. e. in (2005) studied extends earlier results on bias – corrected estimators for the fixed effects dynamic panel data model,[3]. Gorgens a. e. in (2008)discussed efficient estimation of nonlinear dynamic panel data models with application to smooth transition models ,they explores estimation of a class of nonlinear dynamic panel data models with additive unobserved individual specific effects,[9]. Feng a. e. in (2015) proposed a panel data Semiparametric varying coefficient model in which covariates (variables affecting the coefficients) are purely categorical,[6]. Ashley and Sun in (2016) proposed subset continuous updating GMM estimators for dynamic panel data models,[1].

The cornerstone of Bayesian methodology is the Bayes theorem, which is known as the principle of inverse probability. It helps us make probability statements about parameters after the sample has been taken. The conditional distribution of the parameters after observing the data is the posterior distribution that summarizes the prior and the sample information posterior information is proportional to sample information times prior knowledge. Attainment of the posterior is only the beginning of the research methodology since the statistical inference will be based on the posterior and predictive distributions that are derived using the Bayes’ rule. However, to obtain the posterior we do need the data and the prior distribution. The choice of the prior distribution depends on the knowledge of the investigator as well as his willingness to incorporate beliefs and theoretical postulates into the methodology. There are explicit rules for selecting prior distributions whether an informative or non-informative prior is preferred. In the spirit of the simplicity postulate, it is reasonable to begin with a simple case, a regression model with a constant term and a regressor. Markov Chain Monte Carlo (MCMC) is both feasible and provides sufficiently accurate results if used with care. Based on the problem at hand, the investigator can utilize either of these tools or a combination of them can be constructed. This makes it possible to sample from the complicated posterior distributions and/or compute posterior moments or any other inferential summary statistic. Calculation of the marginal posterior functions is an important part of Bayesian analysis, for the usual objective is to make inferences about individual parameters and or provide graphs for those marginal posterior densities. MCMC methods facilitate such investigations,[4],[12],[13],[14].

The linear model subject to a set of linear constraints on the coefficients of the model arises commonly in applied econometrics as well as other scientific applications. Constrained parameter
problems arise in a wide variety of applications, including bioassay, actuarial graduation, ordinal data, response surfaces, reliability development testing and variance component model. Typically the motivating economic model restricts the signs of certain coefficients or of known linear combinations of coefficients. One difficulty with restricted estimation, whether it be from a sampling theory or a Bayesian point of view, is that it assumes an unaltering belief in the prior information implied by the linear constraints. Few researchers are likely to be so stubborn that they would continue to believe in the prior constraints, when faced with strong sample information to the contrary. Given linear constraints Bayesian inference is much simpler than classical inference, but standard Bayesian computational methods become impractical when the posterior probability of the linear constraints (under a diffuse prior) is small,[8],[11].

In this paper, we consider a set of linear constraints on the coefficients of the random panel data model. Furthermore, Bayesian approach based on Markov Chain Monte Carlo (MCMC) is employed to making inferences on the model coefficients subject to this constraints.

2. Panel Data Model

Consider the model:

\[ Y_{it} = \mu + \sum_{j=1}^{K} \beta_{j}X_{jit} + \epsilon_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T, \]  

(1)

where, \( Y_{it} \) the value of response variable for \( i^{th} \) unit at time \( t \), \( X_{jit} \) the explanatory variables, \( \mu, \beta_{j}, \) \( j = 1, ..., K \) are fixed parameters and \( \epsilon_{it} \) is an error term with \( \epsilon_{it} \sim N(0, \sigma_{\epsilon}^2) \).

Now, if the parameter \( \mu \) is specified as:

\[ \mu = \beta_{0} + u_{i}, \]  

(2)

where, \( u_{i} \sim N(0, \sigma_{u}^2) \), then, the model (1) is

\[ Y_{it} = \beta_{0} + \sum_{j=1}^{K} \beta_{j}X_{jit} + u_{i} + \epsilon_{it}. \]  

(3)

The model (3) is rewrite as follows:

\[ Y_{it} = \beta_{0} + \sum_{j=1}^{K} \beta_{j}X_{jit} + \omega_{it}, \]  

(4)

where, \( \omega_{it} = u_{i} + \epsilon_{it} \), \( \omega_{it} \sim N(0, \sigma_{\omega}^2), \sigma_{\omega}^2 = \sigma_{\epsilon}^2 + \sigma_{u}^2 \), thus by using matrix notation the model (4) is

\[ Y = \Phi \theta + \omega, \]  

(5)

where, \( \Phi = [e, X] \), \( e = [1, 1, ..., 1]^T \) has length \( NT \), \( Y = [Y_{11}, ..., Y_{1T}, Y_{21}, ..., Y_{2T}, ..., Y_{N1}, ..., Y_{NT}] \) is a \( NT \times K \) design matrix of fixed effects, \( \theta = [\beta_{0}, \beta_{1}, ..., \beta_{K}]^T \) has length \( K + 1 \), and \( \omega = [\omega_{11}, ..., \omega_{1T}, \omega_{21}, ..., \omega_{2T}, ..., \omega_{N1}, ..., \omega_{NT}]^T \) has length \( NT \). From model (5), we have \( Y \sim N(\Phi \theta, \Psi) \), where

\[ \Psi = E(\omega \omega^T) = I_{N} \otimes (\sigma_{\epsilon}^2 I_{T} + \sigma_{u}^2 ee^T) = \sigma_{\epsilon}^2 (I_{N} \otimes I_{T}) + \sigma_{u}^2 (I_{N} \otimes ee^T), \]

replace \( I_{T} \) by \( (E_{T} + J_{T}) \) and \( ee^T \) by \( T J_{T} \), where \( J_{T} = \frac{1}{T} ee^T \) and \( E_{T} = I_{T} - J_{T} \), then

\[ \Psi = \sigma_{\epsilon}^2 [I_{N} \otimes (E_{T} + J_{T})] + \sigma_{u}^2 (I_{N} \otimes T J_{T}) = \sigma_{\epsilon}^2 (I_{N} \otimes E_{T}) + \sigma_{u}^2 (I_{N} \otimes J_{T}) + T \sigma_{\epsilon}^2 (I_{N} \otimes J_{T}), \]

by collecting terms with the same matrices, we get

\[ \Psi = \sigma_{\epsilon}^2 (I_{N} \otimes E_{T}) + (\sigma_{\epsilon}^2 + T \sigma_{\epsilon}^2) (I_{N} \otimes J_{T}) = \sigma_{\epsilon}^2 Q + \sigma_{u}^2 P, \]

where, \( \sigma_{\epsilon}^2 = (\sigma_{\epsilon}^2 + T \sigma_{\epsilon}^2) \) and \( \Psi^{-1} = \frac{Q}{\sigma_{\epsilon}^2} + \frac{P}{\sigma_{u}^2} \), \( |\Psi| \) is product of its characteristic roots, [2], [3], [4].

3. Constrained panel data model

In this section we consider a set of linear constraints on the coefficients of the random panel data model (5). Furthermore, it investigates the inferences on the random panel data model. Consider the model (5) above, we assume that
\( R\theta = r, \) \hspace{1cm} (6)

where \( R \) is \( m \times K^* \), \( r \) is \( m \times 1 \), \( \theta \) is \( K^* \times 1 \) and \( K^* = K+1 \).

The likelihood function is

\[
L(\theta | \sigma^2_{\epsilon}, \sigma^2) = (2\pi)^{-NT} |\Psi^{-1}|^{-1} \exp \left[ (Y - F\theta)^T \Psi^{-1} (Y - F\theta) \right].
\]

Therefore, it follows that

\[
\hat{\theta} = (F^T \Psi^{-1} F)^{-1} (F^T \Psi^{-1} Y), \quad \text{with } E(\hat{\theta}) = \theta, \quad \text{and } \text{Var}(\hat{\theta}) = (F^T \Psi^{-1} F)^{-1},
\]

\[
\sigma^2 = \frac{1}{N(T-1)} (Y - F\hat{\theta})^T Q (Y - F\hat{\theta}),
\]

\[
\sigma_1^2 = \frac{1}{N} (Y - F\hat{\theta})^T P (Y - F\hat{\theta}).
\]

Now, from above, we can see

\[
Y - F\theta = Y - F\hat{\theta} + F\hat{\theta} - F\theta = Y - \hat{Y} + F\hat{\theta} - F\theta \rightarrow Y - F\theta = \hat{\omega} + F\hat{\theta} - F\theta ,
\]

where

\[
\hat{\theta} = Y - \hat{\omega} \text{ and } \hat{\omega} = Y - \hat{\theta},
\]

\[
(Y - F\theta)^T \Psi^{-1} (Y - F\theta) = (Y - F\hat{\theta} + F\hat{\theta} - F\theta)^T \Psi^{-1} (Y - F\hat{\theta} + F\hat{\theta} - F\theta)
\]

\[
= \left( \hat{\omega} + F(\hat{\theta} - \theta) \right)^T \Psi^{-1} \left( \hat{\omega} + F(\hat{\theta} - \theta) \right)
\]

\[
\rightarrow (Y - F\theta)^T \Psi^{-1} (Y - F\theta) = \hat{\omega}^T \Psi^{-1} \hat{\omega} + \hat{\omega}^T \Psi^{-1} F(\hat{\theta} - \theta) + (\hat{\theta} - \theta)^T F^T \Psi^{-1} F(\hat{\theta} - \theta).
\]

Since, \( \hat{\theta} = (F^T \Psi^{-1} F)^{-1} (F^T \Psi^{-1} Y) \rightarrow \hat{\theta} = F\hat{\theta} = F(F^T \Psi^{-1} F)^{-1} (F^T \Psi^{-1} Y) \)

\[
\rightarrow \hat{\omega} = Y - \hat{\theta} = Y - F(F^T \Psi^{-1} F)^{-1} (F^T \Psi^{-1} Y) = [I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] Y.
\]

In the following theorem we establish some properties of the matrix \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] \)

**Theorem 1:** The matrix \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] \)

(i) \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] \) \(\neq \) \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}]^T \)

(ii) \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] F = 0 \)

(iii) \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] = 0 \)

(iv) \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] \) idempotent of rank \((NT - K - 1)\).

**Proof:**

(i) Since \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] = [I - \Psi^{-1} F(F^T \Psi^{-1} F)^{-1} F^T] \),

(ii) \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] F = [I - \Psi^{-1} F(F^T \Psi^{-1} F)^{-1} F^T] \Psi^{-1} F \)

\[
= \Psi^{-1} F - \Psi^{-1} F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} F = 0.
\]

(iii) \(F^T \Psi^{-1} [I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] = F^T \Psi^{-1} [I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] \)

\[
= F^T \Psi^{-1} - F^T \Psi^{-1} F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} = 0.
\]

(iv) \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}]^2 = [I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] \times \)

\[
[I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] = I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} \]

\[
+ F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} = I - 2 F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} \]

and the rank of the matrix \([I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] = \text{tr}[I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] \)

\[
= [\text{tr} I - \text{tr}(F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1})] = [NT - \text{tr}(F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1})] = NT - K - 1.
\]

Now, by theorem (1) above the second term of equation (7) is
\[ \hat{\omega}^T \Psi^{-1} F(\tilde{\theta} - \theta) = ([I - F(T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] Y)^T \Psi^{-1} F(\tilde{\theta} - \theta) \]
\[ = Y^T [I - F(T \Psi^{-1} F)^{-1} F^T \Psi^{-1}]^T \Psi^{-1} F(\tilde{\theta} - \theta) \]
\[ = Y^T.0.(\tilde{\theta} - \theta) = 0, \]
and the third term of equation (7)
\[ (\tilde{\theta} - \theta)^T F^T \Psi^{-1} \hat{\omega} = (\tilde{\theta} - \theta)^T F^T \Psi^{-1} \cdot [I - F(T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] Y \]
\[ = (\tilde{\theta} - \theta)^T.0.Y = 0. \]
Therefore, the first term of equation (7) is
\[ \hat{\omega}^T \Psi^{-1} \hat{\omega} = (Y - F \tilde{\theta})^T \frac{Q}{\sigma_1^2} + \frac{P}{\sigma_1^2} (Y - F \tilde{\theta}) = - \frac{N(T - 1) \beta_2^2}{2 \sigma_2^2} - \frac{N \beta_2^2}{2 \sigma_2^2}. \]
Then, the equation (7) is
\[ (Y - F \tilde{\theta})^T \Psi^{-1} (Y - F \tilde{\theta}) = \hat{\omega}^T \Psi^{-1} \hat{\omega} + (\tilde{\theta} - \theta)^T F^T \Psi^{-1} F(\tilde{\theta} - \theta). \]
\[ = - \frac{N(T - 1) \beta_2^2}{2 \sigma_2^2} + \frac{N \beta_2^2}{2 \sigma_2^2} + (\tilde{\theta} - \theta)^T F^T \Psi^{-1} F(\tilde{\theta} - \theta). \]
Thus, the likelihood function is
\[ L(Y|\theta, \sigma_2^2, \sigma_2^2) = (2 \pi)^{-NT/2} (\sigma_2^2)^{-N} (\sigma_2^2)^{-N(T - 1)/2} \left( -\frac{1}{2} (Y - F \tilde{\theta})^T \frac{Q}{\sigma_1^2} + \frac{P}{\sigma_1^2} (Y - F \tilde{\theta}) \right) \times \]
\[ \exp\left( -\frac{1}{2} (\tilde{\theta} - \theta)^T \Psi^{-1} F(\tilde{\theta} - \theta) \right) \]
\[ \rightarrow L(Y|\theta, \sigma_2^2, \sigma_2^2) = (2 \pi)^{-NT/2} (\sigma_2^2)^{-N} \exp\left\{ -\frac{1}{2 \sigma_2^2} (Y - F \tilde{\theta})^T P (Y - F \tilde{\theta}) \right\} (\sigma_2^2)^{-N(T - 1)/2} \]
\[ \times \exp\left\{ -\frac{1}{2} \exp\left\{ \left( \tilde{\theta} - \theta \right)^T F^T \Psi^{-1} F \left( \tilde{\theta} - \theta \right) \right\} \right\} \]
\[ \rightarrow L(Y|\theta, \sigma_2^2, \sigma_2^2) = (2 \pi)^{-NT/2} (\sigma_2^2)^{-N} \exp\left\{ -\frac{N \beta_2^2}{2 \sigma_2^2} \right\} \times \exp\left\{ -\frac{N(T - 1) \beta_2^2}{2 \sigma_2^2} \right\} \]
\[ \times \exp\left\{ -\frac{1}{2} (\tilde{\theta} - \theta)^T F^T \Psi^{-1} F (\tilde{\theta} - \theta) \right\} \]
\[ \rightarrow L(Y|\theta, \sigma_2^2, \sigma_2^2) = (2 \pi)^{-NT/2} (\sigma_2^2)^{-N} \exp\left\{ -\frac{N \beta_2^2}{2 \sigma_2^2} \right\} \times \exp\left\{ -\frac{N(T - 1) \beta_2^2}{2 \sigma_2^2} \right\} \times \frac{1}{2} (\tilde{\theta} - \theta)^T F^T \Psi^{-1} F (\tilde{\theta} - \theta) \}
\]
with \( R \theta = r \).

4. Bayesian Constrained panel data model

To specify a complete Bayesian model, we need a prior distribution on \((\theta, \sigma_2^2, \sigma_2^2)\). We will use the uniform distribution \(U(0,1)\) of the vector parameters \(\theta\), as well as we will assume that the prior distribution on \(\alpha_2^2\) and \(\alpha_2^2\) are inverse gamma with parameters \(\alpha_2, \beta_2\alpha_2\) and \(\beta_2\) respectively. i.e.
\[ \pi_0(\sigma_2^2) = \frac{\beta_2 \alpha_2}{\Gamma(\alpha_2)} (\alpha_2)^{-(\alpha_2+1)} \exp\left\{ - \frac{\beta_2}{\sigma_2^2} \right\} \]
\[ \pi_0(\sigma_2^2) = \frac{\beta_2 \alpha_2}{\Gamma(\alpha_2)} (\alpha_2)^{-(\alpha_2+1)} \exp\left\{ - \frac{\beta_2}{\sigma_2^2} \right\}, \text{where } \alpha_2, \beta_2, \alpha_1, \alpha_1, \beta_1 \text{ are hyperparameters that determine the priors and must be chosen by the statistician}. \]
\[ \rightarrow \pi_0(\theta, \sigma_2^2, \sigma_2^2) \propto \frac{\beta_2 \alpha_2}{\Gamma(\alpha_2)} (\alpha_2)^{-(\alpha_2+1)} \beta_2 \alpha_2 \beta_1 \alpha_1 \sigma_2^2 \exp\left\{ - \frac{\beta_2}{\sigma_2^2} \right\} \]
\[ \text{From the model (5) we have } \forall \theta, \sigma_2^2, \sigma_2^2 \sim N_{NT}(F \theta, \Psi). \text{ Then, the joint posterior density of the coefficients } \theta \text{ and the variances } \sigma_2^2 \text{ and } \sigma_2^2 \text{ given by the expression:} \]
\[ \pi_1(\theta|\sigma_2^2, \sigma_2^2) \propto L(Y|\sigma_2^2, \sigma_2^2) \times \pi_0(\theta|\sigma_2^2, \sigma_2^2), \text{ with } R \theta = r \]
\[
\pi_1(\theta | \sigma_e^2, \sigma_1^2, Y) \propto (2\pi)^{-NT} \frac{N}{2} \left( \frac{1}{2} \right) \left( \frac{1}{2} \right) \exp \left\{ -\frac{1}{2} \left( \theta - \hat{\theta} \right)^T F^T \psi^{-1} F \left( \theta - \hat{\theta} \right) \right\} \times \\
\beta_1 \frac{N}{2} \left( \frac{1}{2} \right)^{\frac{N}{2}} \left( \frac{1}{2} \right)^{\frac{N}{2}} \left( \frac{1}{2} \right)^{\frac{N}{2}} \left( \frac{1}{2} \right)^{\frac{N}{2}} \exp \left\{ -\frac{1}{2} \left( \theta - \hat{\theta} \right)^T F^T \psi^{-1} F \left( \theta - \hat{\theta} \right) \right\}
\]

From (9), we can deduce the following conditional and marginal posterior distributions
\[
\pi_1(\theta | \sigma_e^2, \sigma_1^2, Y) \propto \exp \left\{ -\frac{1}{2} \left( \theta - \hat{\theta} \right)^T F^T \psi^{-1} F \left( \theta - \hat{\theta} \right) \right\},
\]

and
\[
\pi_1(\sigma_e^2 | \theta, \sigma_1^2, Y) \propto (\sigma_e^2)^{-\left( a_e + \frac{N(T-1)}{2} \right)} \left( \sigma_e^2 \right)^{\frac{N}{2}} \exp \left\{ -\frac{1}{2} \left( \theta - \hat{\theta} \right)^T F^T \psi^{-1} F \left( \theta - \hat{\theta} \right) \right\},
\]

Therefore, it follows that
\[
\left( \theta | \sigma_e^2, \sigma_1^2, Y \right) \sim \mathcal{N} \left( \hat{\theta}, (F^T \psi^{-1} F)^{-1} \right),
\]

\[
\sigma_e^2 | \theta, \sigma_1^2, Y \sim \text{IG} \left( \alpha_e + \frac{N(T-1)}{2}, \beta_e + \frac{1}{2} N(T-1) \right).
\]

Now, let A is non-singular matrix such that \( A(F^T \psi^{-1} F)^{-1} A^T = I \), and \( \eta = A \theta \) subject to the constraint \( R \theta = r \), then \( RA^{-1} A \theta = r \rightarrow R A^{-1} \eta = r \). From (13) we can obtain
\[
E(\eta) = E(A \theta) = A E(\theta) = A \hat{\theta}, \text{ and } \text{Var}(\eta) = \text{Var}(A \theta) = AVa r(\theta) A^T = A(F^T \psi^{-1} F)^{-1} A^T = I
\]

i.e.
\[
E(\eta | \sigma_e^2, \sigma_1^2, Y) \sim \mathcal{N} \left[ A \hat{\theta}, I \right] \text{ subject to } R A^{-1} \eta = r.
\]

5. Data results

In this section, we illustrate our methodology for model (5) with the data set from gross fixed capital formation and gross domestic product by economic activities for public sector at current prices for the years (2005-2015) (Million I.D.) in Iraq for building and construction and wholesale, retail trade, hotels and others where, which were significance.
The number of iterations of the Gibbs sampler used in this case was (10000) iterations, the probability transition matrix for the model (5) with constraints (6) and the posterior density of this matrix as follows

\[ \hat{P} = \begin{pmatrix} 0.649735493803643 & 0.350264506196357 \\ 0.499917665481466 & 0.500082334518534 \end{pmatrix}. \]

Figure (1) shows the posterior density of the transition matrix.

6. Conclusion

The conclusions which are obtained throughout this paper are given as follows:

1. The matrix \( I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} \) have the following properties:

   (i) \( I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} \neq [I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}]^T \),

   (ii) \( [I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}]^T \Psi^{-1} F = 0 \),

   (iii) \( F^T \Psi^{-1} [I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1}] = 0 \).

   (iv) \( I - F(F^T \Psi^{-1} F)^{-1} F^T \Psi^{-1} \) is idempotent of rank \( NT - K - 1 \).
2. Marginal posterior distributions of $\theta$, $\sigma^2_\varepsilon$ and $\sigma^2_1$ are respectively:

$$
(\theta|\sigma^2_\varepsilon, \sigma^2_1, Y) \sim N(\hat{\theta}, (F^T Y^{-1} F)^{-1})
$$

$$(\sigma^2_\varepsilon | \theta, \sigma^2_1, Y) \sim IG\left(\alpha_\varepsilon + \frac{N(T-1)}{2}, \beta_\varepsilon + \frac{1}{2} N(T-1) \hat{\sigma}^2_\varepsilon\right)
$$

$$(\sigma^2_1 | \theta, \sigma^2_\varepsilon, Y) \sim IG(\alpha_1 + \frac{N}{2}, \beta_1 + \frac{1}{2} N \hat{\sigma}^2_1)
$$

3. The posterior of $\eta$, where $\eta = A\theta$ is

$$(\eta|\sigma^2_\varepsilon, \sigma^2_1, Y) \sim N(A\hat{\theta}, I).
$$

4. The probability transition matrix for the model (5) with constraints (6) as follows

$$
\hat{P} = \begin{pmatrix}
0.649735493803643 & 0.350264506196357 \\
0.499917665481466 & 0.500082334518534
\end{pmatrix}
$$

References

[1] Ashley, Richard A, Sun, Xiaojin," subset continuous updating GMM estimators for dynamic panel data model, Econometrics,4,47;doi:10.3390/econometrics 4040047,(2016).

[2] Baltagi, badi, " Econometric Analysis of panel data " , John Wily & Sons Inc. third edition,(2005).

[3] Bun, Maurice J.G. and Carree, Martin A., "Bias corrected estimation in dynamic panel data model with heteroscedasticity ",UVA Econometrics, (2005).

[4] Congdon, Peter," Bayesian Statistical Modelling", second edition, Wiley Jon & Sons, Ltd, (2006).

[5] Elhorst,J. Paul, "Panel data models extended to spatial error autocorrelation or spatially lagged dependent variable "university of Groningen, university medical center Groningen, (2001).

[6] Feng, Gao, Jiti , Peng Bin and Zhang, Xiaohui," A varying coefficient panel data model with fixed effects ",ISSN 1440-771X, Monash University, (2015).

[7] Frees, Edward W. "Longitudinal and Panel Data: Analysis and Applications for the Social Sciences", This draft is partially funded by the Fortis Health Insurance Professorship of Actuarial Science. Cambridge University Press, (2004).

[8] Geweke, John ," Bayesian Inference for Linear Models Subject to Linear Inequality Constraints", this work was supported in part by National Science Foundation Grant SES 9210070, (1995).

[9] Gorgens, Tue, Skeels, Christopher L. and Wurtz, Allan H., "efficient estimation of nonlinear dynamic panel data models with application to smooth transition models", (2008).

[10] Hurlin, Christophe, "Testing granger causality in heterogeneous panel data model with fixed coefficients", university Paris, (2004).

[11] Lewis, John R.," Bayesian Restricted Likelihood Method ",Technical Report No. 878, (2014).

[12] Marin, Jean-Michel, Robert, Christian P., "Bayesian Essentials With R", second edition, Springer, (2014).

[13] Ruppert, David ,Wand, M. P., and Carroll, R. J., "Semiparametric Regression", Cambridge University Press, (2003).

[14] Wakefield, Jon," Bayesian and Frequentist Regression Methods", Springer, (2013).