MONASH University

MONASH BUSINESS SCHOOL

ISSN 1440-771X

Department of Econometrics and Business Statistics

http://business.monash.edu/econometrics-and-business-statistics/research/publications

Regime Switching Panel Data Models with Interative Fixed Effects

Tingting Cheng, Jiti Gao and Yayi Yan

November 2018

Working Paper 21/18

AACSB ACCREDITED EFMD EQUIS ACCREDITED ASSOCIATION OF MBAS AMBA ACCREDITED

Regime Switching Panel Data Models with Interactive Fixed Effects

Tingting Cheng[‡1], Jiti Gao[*,2] and Yayi Yan[‡]

School of Finance, Nankai University[‡],

Department of Econometrics and Business Statistics, Monash University[*]

Abstract: In this paper, we introduce a regime switching panel data model with interactive fixed effects. We propose a maximum likelihood estimation method and develop an expectation and conditional maximization algorithm to estimate the unknown parameters. Simulation results show that the algorithm works well in finite samples. The biases of the maximum likelihood estimators are negligible and the root mean squared errors of the maximum likelihood estimators decrease with the increase of either cross-sectional units N or time periods T.

Keywords: ECM algorithm; Interactive effect; Maximum likelihood Estimation; Panel data model; Regime switching

JEL classifications: C23, C32

1 Introduction

Recently, there has been a growing literature on panel data models with interactive fixed effects (e.g. Bai 2009; Bai and Li 2014; Moon and Weidner 2015, 2017), where the individual fixed effects, called factor loadings, interact with common time specific effects, called factors. These models are very flexible since it allows for the factors to affect each individual with a different loading, and the conventional additional fixed effects and time-specific effects arise as a special case. These models have been widely studied in various economics disciplines, for example in asset pricing (e.g. Gagliardini et al. 2016), empirical macro (e.g. Gobillon and Magnac 2016), forecasting (e.g. Gonçalves et al. 2017), and empirical labor economics (e.g. Hagedorn et al. 2015). The major advantage of the panel data model with interactive fixed effects is that it provides researchers with a tractable way to capture the co-movement of many individual economic series and model

[1] *Corresponding author*: School of Finance, Nankai University, China. Email: chengtingting2015@outlook.com.
The first and the third authors were supported financially by the National Natural Science Foundation of China (Project No. 71803091) and by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (Project No. 18YJC790015).

[2] The second author acknowledges the support of the Australian Research Council Discovery Grants Program under Grant: DP150101012 & DP170104421.

cross-sectional dependence over individual units. However, it does not uncover possible structural changes over time. Such dramatic changes are, indeed, quite common in many areas in economics and finance, and may result from events, such as wars, financial panic, or significant changes in government policies. There are only few studies in the literature to consider this issue. Li et al. (2016) consider estimation of common structural breaks in panel data models with interactive fixed effects. They introduce a penalised principal component estimation procedure with an adaptive group fused LASSO to detect the multiple structural breaks in the models and a post-LASSO method to estimate the regression coefficients. However, their approach does not include a description of the property or dynamic behaviour of the structural breaks. Therefore, our paper proposes another approach to modelling structural changes in panel data models with interactive fixed effects using a Markov regime switching framework, which regard the change in regime as a random variable and include a description of the probability law governing the change in regime.

Specifically, in this paper, we propose a regime switching panel data model with interactive fixed effects. Such a factor structure has received increasing interest in the analysis of panel data in recent years. However, to the best of our knowledge, there is no study on regime switching panel data models with interactive effects. This paper aims to fill this gap. Here we use business cycles as an example to motivate regime switching panel data models with interactive fixed effects. In an influential study, Burns and Mitchell (1947) proposed two features of the business cycle: the first is the co-movement among individual economic variables; the second is regime switching in the evolution of the business cycle. The factor models (e.g. Bai and Ng 2002; Bai 2003; Bai and Li 2012) and the regime switching models (e.g. Hamilton 1989; Diebold et al. 1994; Kim et al. 2008; Chang et al. 2017) have been used extensively by econometricians to capture the co-movement and regime switching aspects of the business cycle in isolation of each other. In order to encompass the two key features of the business cycle, Diebold and Rudebusch (1994) and Kim and Nelson (1998) proposed a multivariate dynamic factor model with regime switching. Another popular approach to capturing the two features of a business cycle is the vector autoregressive model with regime switching (e.g. Krolzig 2013). Our model can also capture nonlinearity and co-movement simultaneously and include the traditional regime switching factor model as a special case.

For this new model, we provide a maximum likelihood estimation method, along with the proposal of an expectation and conditional maximization (ECM) algorithm, to estimate the unknown parameters. Simulation results show that the algorithm works well in finite samples, the biases of the maximum likelihood estimators are negligible and the root mean squared errors of the maximum likelihood estimators decrease with the increase of either cross-sectional units N or time periods T.

The rest of the paper is organized as follows. The model and the maximum likelihood method are proposed in Section 2. An ECM algorithm is provided in Section 3. Section 4 reports the Monte Carlo simulation results. Section 5 concludes the paper.

2 Model

In this section, we first introduce our proposed regime switching panel data model in detail and then present a maximum likelihood estimation method.

2.1 Model specification

In this paper, we consider the following panel data model with N cross-sectional units and T time periods

$$\begin{aligned} y_{it} &= \alpha_i + X_{it1}\beta_1^{s_t} + X_{it2}\beta_2^{s_t} + ... + X_{itK}\beta_K^{s_t} + \lambda_i' f_t + e_{it}, \\ X_{itk} &= \mu_{ik} + \gamma_{ik}' f_t + v_{itk}, \quad k = 1, 2, ..., K, \end{aligned} \tag{2.1}$$

where α_i denotes the fixed effect, $X_{it} = (X_{it1}, X_{it2}, ..., X_{itK})'$ are the K dimensional regressors, f_t denotes the unknown common factors and λ_i denotes the corresponding factor loadings. In the model, we assume that both y_{it} and x_{it} are influenced by the common shocks f_t but we can easily relax this assumption by letting the regressors being impacted by some additional factors that do not necessarily affect the y_{it} equation. An alternative view is that some factor loadings in the y_{it} equation are restricted to be zero.

The parameter of interest is $\beta^{s_t} = (\beta_1^{s_t}, ..., \beta_K^{s_t})'$, although estimation for α_i, λ_i, μ_{ik} and γ_{ik} ($k = 1, 2, ..., K$) will also be discussed. By treating α_i, λ_i, μ_{ik} and γ_{ik} as parameters, we also allow arbitrary correlations between (α_i, λ_i) and (μ_{ik}, γ_{ik}). We consider the model with two regimes $s_t \in \{1, 2\}$ for illustration, and this model can easily be extended to an N-regime switching panel data model. The dynamic behavior of s_t is modeled by a *first-order Markov chain* and the transition probabilities are collected in a 2×2 matrix P known as the transition matrix in which the element p_{ij} denotes the probability $p(s_t = j | s_{t-1} = i)$.

We assume that the number of factors r is fixed and known.[3] Let $X_{it} = (X_{it1}, X_{it2}, ..., X_{itK})'$, $\Gamma_{ix} = (\gamma_{i1}, \gamma_{i2}, ..., \gamma_{iK})'$, $v_{itx} = (v_{it1}, v_{it2}, ..., v_{itK})'$ and $\mu_{ix} = (\mu_{i1}, \mu_{i2}, ..., \mu_{iK})'$. Then we have

$$X_{it} = \mu_{ix} + \Gamma_{ix}' f_t + v_{itx}.$$

Let $\Gamma_i = (\lambda_i, \Gamma_{ix})$, $z_{it} = (y_{it}, X_{it})'$, $\epsilon_{it} = (e_{it}, v_{itx}')'$ and $\mu_i = (\alpha_i, \mu_{ix}')'$. It is easy to see that

$$B_{st} z_{it} = \mu_i + \Gamma_i' f_t + \epsilon_{it}$$

where $B_{st} = \begin{pmatrix} 1 & -(\beta^{s_t})' \\ 0 & I_K \end{pmatrix}$. Let $z_t = (z_{1t}', z_{2t}', ..., z_{Nt}')'$, $\Gamma = (\Gamma_1, \Gamma_2, ..., \Gamma_N)'$, $\epsilon_t = (\epsilon_{1t}', \epsilon_{2t}', ..., \epsilon_{Nt}')$ and $\mu = (\mu_1', \mu_2', ..., \mu_N')'$. We then have

$$(I_N \otimes B_{s_t}) z_t = \mu + \Gamma' f_t + \epsilon_t. \tag{2.2}$$

The idiosyncratic errors $\epsilon_{it} = (e_{it}, v_{itx}')'$ are such that: the e_{it} is independent and identically distributed over t and uncorrelated over i with $E(e_{it}) = 0$ for all $i = 1, ..., N$ and $t = 1, ..., T$. Let Σ_{iie} denote the variance of e_{it}. The v_{it} is also independent and identically distributed over t and uncorrelated over i with $E(v_{itx}) = 0$ for all $i = 1, ..., N$ and $t = 1, ..., T$. Let Σ_{iix} denote the variance of v_{itx}, and e_{it} be independent of v_{jsx} for all (i, j, t, s). Let Σ_{ii} denote the variance matrix ϵ_{it}. So we have $\Sigma_{ii} = diag(\Sigma_{iie}, \Sigma_{iix})$, a block-diagonal matrix. In the model, we allow for arbitrary correlations between $v_{itk}, k = 1, ..., K$. Let $\Sigma_{\epsilon\epsilon}$ denote the variance of $\epsilon_t = (\epsilon_{1t}', \epsilon_{2t}', ..., \epsilon_{Nt}')$. Due to the uncorrelatedness of ϵ_{it} over i, we have $\Sigma_{\epsilon\epsilon} = diag(\Sigma_{11}, \Sigma_{22}, ..., \Sigma_{NN})$, a block-diagonal matrix.

It is well known in factor analysis that the factors and loadings can only be identified up to a rotation. The models considered in this paper also inherit the same identification issue. Therefore, we impose the following commonly used restrictions for model (2.1), which we refer to as IC (identification conditions):

(IC1) $\Sigma_{ff} = I_r$; (IC2) $\Gamma' \Sigma_{\epsilon\epsilon}^{-1} \Gamma = D$, where D is a diagonal matrix with its diagonal elements

[3] We can also relax this assumption and estimate the number of factors using the method in Bai and Ng (2002).

distinct and arranged in descending order.

2.2 Maximum likelihood estimation

Our model can be estimated by maximum likelihood estimation. Let $F_t = (z_t', z_{t-1}', ..., z_1')'$ be a vector containing observed data, and $\theta = (\beta^{s_t}, \Gamma, \Sigma_{\epsilon\epsilon}, P)$ be the vector of model parameters. The conditional log-likelihood function for the observed data is constructed as

$$\ell(z_1, ...z_T) = \log f(z_1|\theta) + \sum_{t=2}^{T} \log f(z_t|F_{t-1};\theta). \tag{2.3}$$

The maximum likelihood estimator $\widehat{\theta}$ is given by $\hat{\theta} = \arg\max \ell(z_1, ...z_T)$ as usual.

To avoid estimating f_t, we can write the conditional likelihood function as

$$\log f(z_t|s_t,\theta) = -\frac{N(K+1)}{2}\log 2\pi + \frac{1}{2}\log|\Sigma_{zz}^{-1}| - \frac{1}{2}[(I_N \otimes B_{s_t})z_t - \mu]'\Sigma_{zz}^{-1}[(I_N \otimes B_{s_t})z_t - \mu],$$

where $\Sigma_{zz} = \Gamma\Gamma' + \Sigma_{\epsilon\epsilon}$. Also, we need not estimate μ as long as we can demean the variables z_t. Let $\dot{z}_t = z_t - \frac{1}{T}\sum_{t=1}^{T} z_t$ and ignore the constant term. We obtain the concentrated likelihood function

$$\log f(z_t|s_t,\theta) = \frac{1}{2}\log|\Sigma_{zz}^{-1}| - \frac{1}{2}[(I_N \otimes B_{s_t})\dot{z}_t]'\Sigma_{zz}^{-1}[(I_N \otimes B_{s_t})\dot{z}_t]. \tag{2.4}$$

To obtain the conditional density function $f(z_t|F_{t-1};\theta)$ for each period, we use the *Markov switching filter* in Hamilton (1989). Let $p(s_t = j|F_t,\theta)$ denote the inference about the value of s_t based on data obtained through date t and knowledge of the population parameters θ. Collect these conditional probabilities $p(s_t = j|F_t,\theta)$ for $j = 1,2$ in an (2×1) vector denoted $\widehat{\xi}_{t|t}$. One could also imagine forming forecasts of how likely the process is to be in regime j in period $t+1$ given observations obtained through date t. Collect these forecasts in an (2×1) vector $\widehat{\xi}_{t+1|t}$, which is a vector whose jth element represents $p(s_{t+1} = j|F_t,\theta)$.

The optimal inference and forecast for each date t in the sample can be found by iterating on the following pair of equations:

$$\widehat{\xi}_{t|t} = \frac{(\widehat{\xi}_{t|t-1} \odot \eta_t)}{(\widehat{\xi}_{t|t-1}'\eta_t)}, \quad \widehat{\xi}_{t+1|t} = P \cdot \widehat{\xi}_{t|t}. \tag{2.5}$$

Here η_t represents the (2×1) vector whose jth element is the conditional density in (2.4) and the symbol $\odot$ denotes element-by-element multiplication. Given a starting value $\hat{\xi}_{1|0} = [0.5, 0.5]'$ and an assumed value for the population parameter, we can iterate by (2.5) for $t = 1, 2, ..., T$ to calculate the values of $\hat{\xi}_{t|t}$ and $\hat{\xi}_{t+1|t}$ for each date t in the sample. Then the conditional density function $f(z_t|F_{t-1};\theta)$ for each period can be calculated as

$$f(z_t|F_{t-1};\theta) = \hat{\xi}'_{t|t-1} \cdot \eta_t.$$

The likelihood function can be maximized through numerical methods such as Newton-Raphson when the number of cross-section (N) is not large.

3 An ECM algorithm

To estimate model (2.1) by the maximum likelihood method, we use the expectation and conditional maximization (ECM) algorithm. In this algorithm, we need to make an inference about the unobserved variables, specifically, $E(f_t|z_t,\theta)$, $E(f_t f_t'|z_t,\theta)$, $p(s_t|F_T,\theta)$ and $p(s_t, s_{t-1}|F_T,\theta)$ in the expectation step. The way of estimating these unobserved variables are given as follows.

The joint distribution of z_t and f_t is assumed to be normal as follows:

$$\begin{bmatrix} f_t \\ (I_N \otimes B_{s_t})\dot{z}_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_r & \Gamma' \\ \Gamma & \Sigma_{zz} \end{bmatrix} \right)$$

with $\Sigma_{zz} = \Gamma\Gamma' + \Sigma_{\epsilon\epsilon}$. It follows that the conditional mean is $E(f_t|z_t,\theta) = \Gamma'\Sigma_{zz}^{-1}(I_N \otimes B_{s_t})\dot{z}_t$ and the conditional variance is $Var(f_t|z_t,\theta) = I_r - \Gamma'\Sigma_{zz}^{-1}$. The conditional second moment is given by $E(f_t f_t'|z_t,\theta) = Var(f_t|z_t,\theta) + E(f_t|z_t,\theta)E(f_t'|z_t,\theta)$.

Smoothed inferences can be calculated using an algorithm developed by Kim (1994). We collect $p(s_t|F_T,\theta)$ in $\xi_{t|T}$. This algorithm can be written as

$$\xi_{t|T} = \xi_{t|t} \odot \{P' \cdot [\xi_{t+1|T}(\div)\xi_{t+1|t}]\},$$

$$p(s_t = j, s_{t-1} = i|F_T,\theta) = p(s_t = j|F_T,\theta) \cdot \frac{p_{ij} \cdot p(s_{t-1} = i|F_t,\theta)}{p(s_t = j|F_t,\theta)},$$

where the sign $(\div)$ denotes element-by-element division. The smoothed probabilities $\xi_{t|T}$ are found by iterating backward for $t = T-1, T-2, ..., 1$. This iteration is started with $\xi_{T|T}$, which is obtained from (2.5) for $t = T$. This algorithm is valid only when s_t follows a *first-order Markov chain*, for details, see Kim (1994).

In the CM-steps, we can simply uses $p(s_t|F_T, \theta)$ to reweight the observed data z_t and the expectation of f_t. Calculation of simple sample statistics of ordinary least squares (OLS) and generalized linear squares (GLS) regressions on the weighted data then generate new estimates of θ. In the CM-steps, we also need to update the variance matrix of $\Sigma_{\epsilon\epsilon}$.

The joint distribution of ϵ_t and f_t is also assumed to be normal as follows:

$$\begin{bmatrix} \epsilon_t \\ (I_N \otimes B_{s_t})\dot{z}_t \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_{\epsilon\epsilon} & \Sigma_{\epsilon\epsilon} \\ \Sigma_{\epsilon\epsilon} & \Sigma_{zz} \end{bmatrix}\right).$$

Similarly, we have $Var(\epsilon_t|z_t, \theta) = \Sigma_{\epsilon\epsilon} - \Sigma_{\epsilon\epsilon}(\Gamma\Gamma' + \Sigma_{\epsilon\epsilon})^{-1}\Sigma_{\epsilon\epsilon}$ and the conditional second moments is $E(\epsilon_t\epsilon_t'|z_t, \theta) = Var(\epsilon_t|z_t, \theta) + E(\epsilon_t|z_t, \theta)E(\epsilon_t'|z_t, \theta)$, where $E(\epsilon_t|z_t, \theta)$ is given by

$$E(\epsilon_t|z_t, \theta) = (I_N \otimes B_{s_t})\dot{z}_t - \Gamma E(f_t|z_t, \theta).$$

In the E-step, given the unknown parameters $\theta^{(k)} = \left(\Gamma^{(k)}, \Sigma_{\epsilon\epsilon}^{(k)}, \beta_0^{(k)}, \beta_1^{(k)}, P^{(k)}\right)$, we first calculate the smoothed probabilities $p(s_t|F_T, \theta^{(k)})$ collected in $\widehat{\xi}_{t|T}$, $p(s_t, s_{t-1}|F_T, \theta^{(k)})$ and the conditional optimal forecasts, which are given by:

$$\widehat{\xi}_{t|T} = \widehat{\xi}_{t|t} \odot \{\widehat{P}' \cdot [\widehat{\xi}_{t+1|T}(\div)\widehat{\xi}_{t+1|t}]\},$$

$$p(s_t = j, s_{t-1} = i|F_T, \theta^{(k)}) = p(s_t = j|F_T, \theta^{(k)}) \cdot \frac{p_{ij}^{(k)} \cdot p(s_{t-1} = i|F_t, \theta^{(k)})}{p(s_t = j|F_t, \theta^{(k)})},$$

$$\widehat{f}_t = E(f_t|z_t, \theta^{(k)}) = \Gamma^{(k)\prime}(\Gamma^k\Gamma^{(k)\prime} + \Sigma_{\epsilon\epsilon}^{(k)})^{-1}(I_N \otimes B_{s_t}^{(k)})\dot{z}_t,$$

$$\begin{aligned}\widehat{f_tf_t'} = E(f_tf_t'|z_t, \theta^{(k)}) = & I_r - \Gamma^{(k)\prime}(\Gamma^k\Gamma^{(k)\prime} + \Sigma_{\epsilon\epsilon}^{(k)})^{-1} + \Gamma^{(k)\prime}(\Gamma^k\Gamma^{(k)\prime} + \Sigma_{\epsilon\epsilon}^{(k)})^{-1} \\ & (I_N \otimes B_{s_t}^{(k)})\dot{z}_t\dot{z}_t{}'(I_N \otimes B_{s_t}^{(k)\prime})(\Gamma^k\Gamma^{(k)\prime} + \Sigma_{\epsilon\epsilon}^{(k)})^{-1}\Gamma^{(k)},\end{aligned}$$

$$E(f_tz_t'|z_t, \theta^{(k)}) = \Gamma^{(k)\prime}(\Gamma^k\Gamma^{(k)}\prime + \Sigma_{\epsilon\epsilon}^{(k)})^{-1}(I_N \otimes B_{s_t}^{(k)})\dot{z}_t\dot{z}_t{}'(I_N \otimes B_{s_t}^{(k)\prime}).$$

In the CM-steps, the parameter θ is partitioned into subvectors and we update $\theta^{(k)}$ to $\theta^{(k+1)}$ according to

$$p_{ij}^{(k+1)} = \frac{\sum_{t=2}^{T} p(s_t = j, s_{t-1} = i | F_T, \theta^{(k)})}{\sum_{t=2}^{T} p(s_{t-1} = i | F_T, \theta^{(k)})},$$

$$\Gamma^{\dagger} = \left[\frac{1}{T} \sum_{t=1}^{T} \sum_{s_t=1}^{2} p(s_t | F_T, \theta^{(k)}) E(z_t f_t' | z_t, \theta^{(k)}) \right] \left[\frac{1}{T} \sum_{t=1}^{T} \sum_{s_t=1}^{2} p(s_t | F_T, \theta^{(k)}) E(f_t f_t' | z_t, \theta^{(k)}) \right]^{-1},$$

$$res_t = (I_N \otimes B_{s_t}^{(k)}) \dot{z}_t - \Gamma^{\dagger} E(f_t | z_t, \theta^{(k)}),$$

$$\Sigma_{\epsilon\epsilon}^{\dagger} = T^{-1} \sum_{t=1}^{T} \sum_{s_t=1}^{2} p(s_t | F_T, \theta^{(k)}) res_t res_t' + \Sigma_{\epsilon\epsilon}^{(k)} - \Sigma_{\epsilon\epsilon}^{(k)} (\Gamma^k \Gamma^{(k)\prime} + \Sigma_{\epsilon\epsilon}^{(k)})^{-1} \Sigma_{\epsilon\epsilon}^{(k)},$$

$$\beta_j^{(k+1)} = \left(\sum_{i=1}^{N} \sum_{t=1}^{T} p(s_t = j | F_T, \theta^{(k)}) \dot{x_{it}}' (\Sigma_{ii\epsilon}^{\dagger})^{-1} \dot{x_{it}} \right)^{-1} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} p(s_t = j | F_T, \theta^{(k)}) \dot{x_{it}}' (\Sigma_{ii\epsilon}^{\dagger})^{-1} (\dot{y_{it}} - \lambda_i^{\dagger} E(f_t | z_t, \theta^{(k)})) \right),$$

where $i, j = 1, 2$.

Finally, we need to normalize the updated parameter to meet the identification conditions. The process is given as follows:

$$\Sigma_{\epsilon\epsilon}^{(k+1)} = Dg\{\Sigma_{\epsilon\epsilon}^{\dagger}\},$$

$$\Gamma^{(k+1)} = Q' \Gamma^{\dagger},$$

where Q is an orthogonal matrix and D is a diagonal matrix satisfying $\Gamma^{\dagger} (\Sigma_{\epsilon\epsilon}^{(k+1)})^{-1} \Gamma^{\dagger\prime} = QDQ'$;$Dg\{\cdot\}$ is the operator that sets the entries of it argument to zeros if the counterparts of $E(\epsilon_t \epsilon_t\prime)$ are zeros.

The ECM algorithm works for both $N < T$ and $N > T$ and it is an effective way of estimating our model.

4 Simulation studies

In this section, we conduct a set of simulations to evaluate the performance of our regime switching panel data model and the proposed ECM algorithm.

In this study, simulated data are generated according to

$$\begin{aligned} y_{it} &= X_{it1}\beta_1^{s_t} + X_{it2}\beta_2^{s_t} + \lambda_i' f_t + e_{it}, \\ X_{itk} &= \gamma_{ik}' f_t + v_{itk}, \quad k = 1, 2. \end{aligned} \tag{4.1}$$

The dimensions of f_t are chosen to be 1 and 3. We set $\beta^1 = (\beta_1^1, \beta_2^1)' = (-1, -2)'$ and $\beta^2 = (\beta_1^2, \beta_2^2)' = (1, 2)'$, and f_t, λ_i and γ_{ik} are generated from standard normal distributions. We consider two types of the transition matrix given by (p_{11}, p_{22})=$(0.7, 0.7)$ and $(0.9, 0.9)$, which are respectively corresponding to different persistence of regime switching, and the unobserved states are generated according to the transition matrix. To generate the error term $\epsilon_{it} = (e_{it}, v_{itx}')'$, we first generate an $N(K+1) \times N(k+1)$ standard normal random matrix and sets some elements to zeros so as to get a block diagonal matrix, then we generate an $N(K+1) \times 1$ standard normal random vector. The error term is generated by the vector multiplied by the block diagonal matrix. Once ϵ_{it} and the true states are obtained, we use $z_t = (I_N \otimes B_{s_t})^{-1}\Gamma' f_t + \epsilon_t$ to yield the observable data.

For model (4.1), we consider seven cases of (N, T) which are $(N, T) = (25, 100)$, $(50, 100)$, $(75, 100)$, $(100, 25)$, $(100, 50)$, $(100, 75)$ and $(100, 100)$. For each pair of (N, T), we conduct 200 replications. In addition, we measure the accuracy of the estimates by computing the bias and the root mean squared error (RMSE) of the 200 maximum likelihood (ML) estimates, which are given by

$$Bias(\hat{\theta}) = \frac{1}{200}\sum_{i=1}^{200}(\hat{\theta}_i - \theta_0), \;\; RMSE(\hat{\theta}) = \sqrt{\frac{1}{200}\sum_{i=1}^{200}(\hat{\theta}_i - \theta_0)^2},$$

where $\hat{\theta}_i$ is the estimate of parameter θ_0 for the i-th replication, and θ_0 is the true value of parameters which is $\theta_0 = (\beta_1^1, \beta_2^1, \beta_1^2, \beta_2^2)' = (-1, -2, 1, 2)'$.

Tables 4.1-4.2 report the simulation results based on 200 repetitions. Each cell contains the bias or the RMSE of the 200 ML estimates for the parameter listed in the column heading. From Tables 4.1-4.2, we can obtain the following results:

- For each example, the biases of the ML estimators are very small and negligible.
- For each example, the RMSEs of the ML estimators become smaller with the increase of either N or T.

- The performance of the ML estimators, under the case of different persistence of regime switching, are quite similar.

- The performance of the ML estimators, under the case of different factor numbers, are quite similar.

Table 4.1: The performance of ML estimators in the proposed model (4.1) when $p_{11} = 0.7, p_{22} = 0.7$

N	T	$\beta_1^1 = -1$		$\beta_2^1 = -2$		$\beta_1^2 = 1$		$\beta_2^2 = 2$	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
				$r = 1$					
25	100	0.0003	0.0048	0.0003	0.0044	0.0003	0.0045	-0.0004	0.0050
50	100	0.0000	0.0021	-0.0002	0.0023	0.0002	0.0020	0.0000	0.0023
75	100	0.0001	0.0017	-0.0001	0.0014	0.0001	0.0014	-0.0002	0.0014
100	100	-0.0001	0.0010	0.0001	0.0011	-0.0001	0.0012	-0.0001	0.0011
100	25	0.0001	0.0024	0.0006	0.0033	0.0000	0.0024	0.0000	0.0028
100	50	0.0000	0.0016	0.0002	0.0016	0.0000	0.0017	0.0001	0.0018
100	75	0.0001	0.0013	0.0000	0.0012	0.0000	0.0012	0.0001	0.0013
				$r = 3$					
25	100	0.0001	0.0050	-0.0001	0.0043	0.0002	0.0050	-0.0001	0.0045
50	100	0.0001	0.0025	0.0002	0.0024	0.0000	0.0025	0.0000	0.0022
75	100	-0.0001	0.0015	0.0000	0.0015	-0.0001	0.0014	0.0001	0.0015
100	100	0.0001	0.0012	-0.0001	0.0013	-0.0001	0.0011	0.0000	0.0011
100	25	-0.0001	0.0023	0.0001	0.0024	0.0000	0.0025	-0.0002	0.0027
100	50	0.0000	0.0016	0.0001	0.0016	0.0000	0.0016	0.0001	0.0017
100	75	-0.0001	0.0013	-0.0001	0.0014	0.0001	0.0013	0.0000	0.0014

5 Conclusion

In this paper, we have proposed a *regime switching panel data model with interactive fixed effects*, which substantially generalizes the existing work which either considers panel data models with interactive fixed effects but no regime switching (e.g. Bai 2009; Bai and Li 2014), or panel data models with regime switching but under cross-sectional independence (e.g. Chen 2007). We have proposed a maximum likelihood estimation method and developed an ECM algorithm to estimate the unknown parameters. Our simulation studies have shown that the proposed model and the estimation method have good finite–sample performance.

Table 4.2: The performance of ML estimators in the proposed model (4.1) when $p_{11} = 0.9, p_{22} = 0.9$

N	T	$\beta_1^1 = -1$		$\beta_2^1 = -2$		$\beta_1^2 = 1$		$\beta_2^2 = 2$	
		Bias	RMSE	Bias	RMSE	Bias	RMSE	Bias	RMSE
					$r = 1$				
25	100	0.0001	0.0043	-0.0003	0.0045	-0.0003	0.0050	0.0002	0.0055
50	100	0.0004	0.0021	-0.0001	0.0024	0.0001	0.0021	0.0000	0.0022
75	100	0.0001	0.0015	0.0000	0.0014	0.0001	0.0015	-0.0001	0.0018
100	100	-0.0001	0.0012	-0.0001	0.0011	-0.0001	0.0012	0.0001	0.0011
100	25	-0.0005	0.0032	-0.0003	0.0041	-0.0003	0.0029	0.0002	0.0026
100	50	0.0000	0.0022	0.0000	0.0018	-0.0001	0.0016	0.0001	0.0017
100	75	0.0000	0.0013	0.0000	0.0013	0.0000	0.0014	0.0001	0.0015
					$r = 3$				
25	100	0.0002	0.0044	-0.0001	0.0056	-0.0001	0.0055	-0.0006	0.0047
50	100	0.0002	0.0025	0.0002	0.0026	-0.0001	0.0025	0.0002	0.0021
75	100	0.0001	0.0016	0.0001	0.0015	0.0001	0.0015	-0.0001	0.0015
100	100	0.0001	0.0012	0.0001	0.0011	0.0000	0.0011	0.0000	0.0013
100	25	-0.0002	0.0033	0.0002	0.0035	0.0000	0.0029	-0.0001	0.0026
100	50	0.0000	0.0021	-0.0001	0.0018	-0.0001	0.0018	-0.0001	0.0020
100	75	-0.0001	0.0013	-0.0002	0.0015	0.0000	0.0013	0.0000	0.0016

References

J. Bai. Inferential theory for factor models of large dimensions. *Econometrica*, 71(1):135–171, 2003.

J. Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, 2009.

J. Bai and K. Li. Statistical analysis of factor models of high dimension. *The Annals of Statistics*, 40(1):436–465, 2012.

J. Bai and K. Li. Theory and methods of panel data models with interactive effects. *The Annals of Statistics*, 42(1): 142–170, 2014.

J. Bai and S. Ng. Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221, 2002.

A.F. Burns and W.C. Mitchell. Measuring business cycles. 1947.

Y. Chang, Y. Choi, and J.Y. Park. A new approach to model regime switching. *Journal of Econometrics*, 196(1): 127–143, 2017.

S.W. Chen. Measuring business cycle turning points in japan with the markov switching panel model. *Mathematics and Computers in Simulation*, 76(4):263–270, 2007.

F.X. Diebold and G.D. Rudebusch. Measuring business cycles: A modern perspective. Technical report, National Bureau of Economic Research, 1994.

F.X. Diebold, J.H. Lee, and G.C. Weinbach. Regime switching with time-varying transition probabilities. *Business Cycles: Durations, Dynamics, and Forecasting*, 1:144–165, 1994.

P. Gagliardini, E. Ossola, and O. Scaillet. Time-varying risk premium in large cross-sectional equity data sets. *Econometrica*, 84(3):985–1046, 2016.

L. Gobillon and T. Magnac. Regional policy evaluation: Interactive fixed effects and synthetic controls. *Review of Economics and Statistics*, 98(3):535–551, 2016.

S. Gonçalves, B. Perron, and A. Djogbenou. Bootstrap prediction intervals for factor models. *Journal of Business & Economic Statistics*, 35(1):53–69, 2017.

M. Hagedorn, I. Manovskii, and K. Mitman. The impact of unemployment benefit extensions on employment: The 2014 employment miracle? Technical report, National Bureau of Economic Research, 2015.

J.D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, pages 357–384, 1989.

C.J. Kim. Dynamic linear models with markov-switching. *Journal of Econometrics*, 60(1-2):1–22, 1994.

C.J. Kim and C.R. Nelson. Business cycle turning points, a new coincident index, and tests of duration dependence based on a dynamic factor model with regime switching. *Review of Economics and Statistics*, 80(2):188–201, 1998.

C.J. Kim, J. Piger, and R. Startz. Estimation of markov regime-switching regression models with endogenous switching. *Journal of Econometrics*, 143(2):263–273, 2008.

H.M. Krolzig. *Markov-switching vector autoregressions: Modelling, statistical inference, and application to business cycle analysis*, volume 454. Springer Science & Business Media, 2013.

D. Li, J. Qian, and L. Su. Panel data models with interactive fixed effects and multiple structural breaks. *Journal of the American Statistical Association*, 111(516):1804–1819, 2016.

H.R. Moon and M. Weidner. Linear regression for panel with unknown number of factors as interactive fixed effects. *Econometrica*, 83(4):1543–1579, 2015.

H.R. Moon and M. Weidner. Dynamic linear panel regression models with interactive fixed effects. *Econometric Theory*, 33(1):158–195, 2017.