Regional Convergence in Japan: A Bayesian Spatial Econometrics Perspective

Kazuhiko Kakamu Hajime Wago

Faculty of Law and Economics, Chiba University

Department of Economics, Kyoto Sangyo University

19th International Conference on Computational Statistics
Paris - France, August 22-27 2010

Outline

Convergence, σ- and β-convergence

- The idea of convergence in economics is the hypothesis that poorer economies' per capita incomes will tend to grow at faster rates than richer economies. As a result, all economies should eventually converge in terms of per capita income.
- Developing countries have the potential to grow at a faster rate than developed countries because diminishing returns aren't as strong as in capital rich countries.
- In the economic growth literature the term "convergence" can have two meanings:
    - "σ-convergence" refers to the catch up effect between countries described above.
    - "β-convergence" refers to countries converging to their own steady state long run growth rate.

Previous Literatures

- Convergence hypothesis is one of the main themes in neoclassical growth theory.
- A lot of researches has developed in theoretical and empirical points of view (see Temple, 1999).
- In Japanese cases:
  - Barro and Sala-i-Martin (1992) compared the β- and σ-convergences using Japanese prefecture and US state data.
  - Togo (2002) and Kakamu and Fukushige (2006) examined Markov transition matrices proposed by Quah (1993) and showed that the Ergodic distributions have two or more peaks, that is, non-normality is observed from the empirical results in Japan.

Summary of Descriptive Statistics

year	mean	variance	min	max	skewness
1986	14.475	0.020	14.249	14.983	0.949
1987	14.555	0.020	14.339	15.051	0.825
1988	14.623	0.022	14.371	15.135	0.767
1989	14.711	0.023	14.453	15.264	0.927
1990	14.773	0.022	14.509	15.312	0.892
1991	14.809	0.023	14.546	15.317	0.771
1992	14.809	0.019	14.556	15.278	0.798
1993	14.814	0.018	14.561	15.296	0.916
1994	14.844	0.016	14.566	15.300	0.773
1995	14.859	0.016	14.581	15.264	0.393
1996	14.894	0.017	14.603	15.281	0.356
1997	14.881	0.017	14.585	15.283	0.349
1998	14.855	0.016	14.596	15.258	0.504
1999	14.851	0.014	14.590	15.248	0.523
2000	14.855	0.015	14.569	15.289	0.589
2001	14.818	0.015	14.537	15.255	0.766
2002	14.803	0.016	14.524	15.222	0.579
2003	14.806	0.018	14.529	15.266	0.590
2004	14.800	0.021	14.502	15.333	0.838

From the Summary Statistics

- We can observe that the variance becomes smaller until 1999 and larger after 1999. $\Longrightarrow$ We can conclude that the σ-convergence is observed until 1999.
- We have to mention that the skewness is greater than zero, that is, it implies that the log per capita income is far from normality. $\Longrightarrow$ To take into account the fact in the model, we consider a two-states (higher and lower income states) normal mixture model.

The Model 1

Let y_{it} be the log per capita income in prefecture i in time t.

$$y_{it} \sim \mathcal{N}(\mu_t, \sigma_t^2), \tag{1}$$

where μ_t and σ_t^2 are mean and variance in time t. In addition, it is not explicitly assumed, but the normality is assumed implicitly. From the estimated σ_t^2 , if it becomes smaller over time, we conclude that the σ-convergence is observed.

The Model 2

However, as is pointed out by Kakamu (2009), if we take into account the spatial interaction in the model, the results of σ-convergence may change.

$$y_{it} \sim \mathcal{N}\left(\rho \sum_{j=1}^{n} w_{ij} y_{jt} + \mu_t, \sigma_t^2\right), \tag{2}$$

where w_{ij} is the i,jth element of the weight matrix and it represents the relationship between i and j. In this paper, we consider a contiguity dummy as the elements of weight matrix (see Anselin, 1988).

The Model 3

To take into account the skewness, we consider two-states normal mixture model. Moreover, we consider a Markov switching model to take into account the time-series structure.

$$y_{it}|s_{i,t-1} = j \sim \mathcal{N}(\mu_{kt}, \sigma_{kt}^2), \tag{3}$$

$$\Pr(s_{it} = k|s_{i,t-1} = j) = p_{kj}, \tag{4}$$

where s_{it} for j, $k = 1, 2$ is defined as a latent state variable (see, e.g. Frühwirth-Schnatter, 2006) and we assume $\mu_2 > \mu_1$ for identification.

The Model 4

We consider the combined model of the above two models as follows:

$$y_{it}|s_{i,t-1} = j \sim \mathcal{N}\left(\rho \sum_{j=1}^{n} w_{ij} y_{jt} + \mu_{kt}, \sigma_{kt}^2\right), \tag{5}$$

$$\Pr(s_{it} = k|s_{i,t-1} = j) = p_{kj}. \tag{6}$$

Prior Distributions

Since we adopt a Bayesian approach, we complete the model by specifying the prior distribution over the parameters. Therefore, we apply the following prior distribution:

$$\pi(\mathbf{\Pi}, \mu, \boldsymbol{\sigma^2}, \rho) = \pi(\mathbf{\Pi}) \prod_{t=1}^{T} \left\{ \prod_{j=1}^{2} \pi(\mu_{jt}) \pi(\sigma_{jt}^2) \right\} \pi(\rho_t),$$

where $\mathbf{\Pi} = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix}$,
$\mu = \{\{\mu_{jt}\}_{j=1}^2\}_{t=1}^T$, $\boldsymbol{\sigma^2} = \{\{\sigma_{jt}^2\}_{j=1}^2\}_{t=1}^T$, and $\rho = \{\rho_t\}_{t=1}^T$.

$$p_{kk} \sim \mathcal{BE}(a_0, b_0),\ \ \mu_{kt} \sim \mathcal{N}(\mu_0, \tau_0^2),$$
$$\sigma_{kt}^2 \sim \mathcal{IG}(\nu_0/2, \lambda_0/2),\ \ \rho_t \sim \mathcal{U}(-1, 1),$$

Joint Posterior Distribution

Given a prior density $\pi(\mathbf{\Pi}, \mu, \sigma^2, \rho)$ and the likelihood function $L(\mathbf{Y}, \mathbf{S}|\mathbf{\Pi}, \mu, \sigma^2, \rho, \mathbf{W})$, the joint posterior distribution can be expressed as

$$\pi(\mathbf{\Pi}, \mu, \sigma^2, \rho|\mathbf{Y}, \mathbf{S}, \mathbf{W}) \propto \pi(\mathbf{\Pi}, \mu, \sigma^2, \rho)L(\mathbf{Y}, \mathbf{S}|\mathbf{\Pi}, \mu, \sigma^2, \rho, \mathbf{W}). \tag{7}$$

Empirical Results

- We use per capita income prepared by Cabinet Office, Government of Japan from 1986 to 2004.
- The weight matrix **W** consists of contiguity dummy variables, proposed by Kakamu et al. (2008).
- For the prior distributions, we set the hyper-parameters as follows:

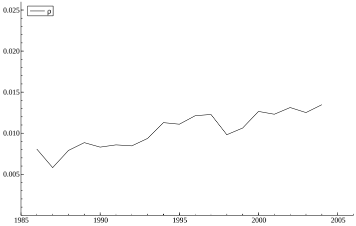
$$a_0 = 0.01,\ b_0 = 0.01,\ \mu_0 = 0.0,\ \tau_0 = 100,\ \nu_0 = 2.0,\ \lambda_0 = 0.01.$$

- We run the MCMC algorithm for $10,000$ iterations after a burn-in phase of $10,000$ iterations.

The trend of ρ

Model 2

Model 4

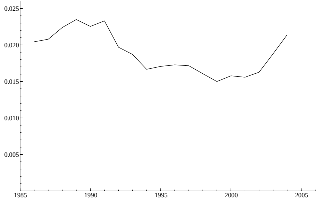


The trend of ρ

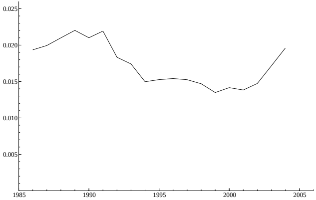
- If we compare the result of Model 2 with Model 4, we can observe that the the spatial interactions of Model 4 is smaller than those of Model 2.
- It implies that the heterogeneity, which is one of the sources of spatial interaction, is captured by two-states normal mixture representations.
- Even if we focus on the results Model 4, the posterior means are slightly increasing and the 95% credible intervals do not include zero, although the magnitude of the spatial interaction is small.
- We can conclude that the spatial interaction plays a weak but important role in examining per capita income in Japan and the role becomes important over time.

The trend of σ^2

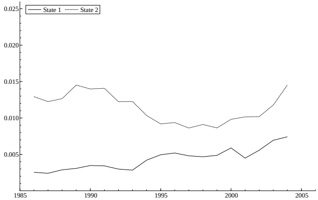
Model 1



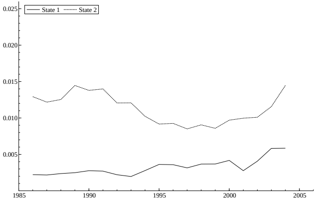
Model 2



Model 3



Model 4



The trend of σ^2

- We can observe that even if we take into account the spatial interaction, the trend of σ^2 does not change.
- As is pointed out by Kakamu (2009), we can also observe that if we consider the spatial interaction, the variances in spatial model become smaller than those of the model without spatial interaction.
- We can conclude that the variance reduction effects are invariant over time and after that we discuss the results from spatial models (Model 2 and 4).

Summary of Empirical Results

- As is pointed out from the descriptive statistics, we can observe that the variances reduce until 1999 and increase after that from the result of Model 2.
- If we move to the result of Model 4, such a σ-convergence is observed only in state 2 (higher income state) until 1997 and the variances of state 1 (lower state income) continue increasing over time.
- We can conclude that the simple σ-convergence model captures the effect of higher income state mainly and there is some possibility that the effect of lower income state.

Conclusions

- As is pointed out from the descriptive statistics, we can observe that the variances reduce until 1999 and increase after that from the result of Model 2.
- If we move to the result of Model 4, such a σ-convergence is observed only in state 2 (higher income state) until 1997 and the variances of state 1 (lower state income) continue increasing over time.
- We can conclude that the simple σ-convergence model captures the effect of higher income state mainly and there is some possibility that the effect of lower income state.
- The trend of the skewness seems to be related to the trend of the variance in state 2. $\Longrightarrow$ We can conclude that the skewness in per capita income in Japan is captured by the two-states normal mixture representation.