Economics Bulletin

Volume 30, Issue 1

Bayesian estimation of spatial externalities using regional production function: the case of China and Japan

Yoshihiro Hashiguchi
Kobe University

Abstract

This paper used regional panel data for Chinese provinces from 1979 to 2003, and for Japanese prefectures from 1955 to 1998, to estimate the spatial externalities (or spatial multiplier effects) using a production function and Bayesian methodology, and to investigate the long-run behavior of the spatial externalities of each country. According to the estimation results, China's spatial externalities increased its domestic production significantly after 1994, which tended to increase until 2003. Before 1993, however, its spatial externalities were not significant. Japan's spatial externalities showed fluctuating values throughout the sample period. Furthermore, the movement of the spatial externalities was correlated with Japan's business conditions: the externalities showed a high value in the economic boom, and a low value in the economic depression. This could mean that spatial externalities correlate mainly with business conditions.

The author wishes to thank the referee for his valuable comments and suggestions.
Citation: Yoshihiro Hashiguchi, (2010) "Bayesian estimation of spatial externalities using regional production function: the case of China and Japan", *Economics Bulletin*, Vol. 30 no.1 pp. 751-764.
Submitted: Oct 16 2009. **Published:** March 18, 2010.

1. Introduction

Over the past few years, concern about spatial externalities in the field of regional economics has risen (Anselin 2003). Spatial externalities are the external effects that spread over several regions, implying that knowledge or ideas that improve the technology of production spill over from one region to the other; thus technical progress in one region brings about an improvement in productivity, not only in its own region, but also in other nearby regions. It seems natural to assume that a regional economy is influenced to some extent by its spatial externalities. To what extent do spatial externalities have an effect on a regional economy? Are their effects only trivial or are they essential for regional economic growth? Therefore, to understand economic growth it is important to measure quantitatively the effect of spatial externalities.

Some studies have attempted an empirical analysis of regional economic growth that takes account of spatial externalities. Ertur and Koch (2007) developed a spatially augmented Solow model, by introducing spatial externalities into the traditional Solow model, and estimated the impact of saving, population growth, and neighborhood on both real income and its growth rate. They used the data from Penn World Tables version 6.1 (91 countries, 1960–1995; Heston et al. 2002) and spatial econometric tools (Anselin 1988; 2001), and concluded that spatial externalities were significant. In other studies, Váya et al. (2004), Fingleton and López-Bazo (2006), Olejnik (2008), and Pfaffermayr (2009) each undertook an econometric analysis of economic growth in Europe, using a version of the spatially augmented Solow model. Each of these studies emphasized the importance of spatial externalities for economic growth.

The problem with the previous studies lies in the fact that few of these studies have attempted to clarify the long-run behavior of spatial externalities. Kakamu et al. (2007) estimated Japan's production function including spatial externalities, and examined year-to-year change in spatial externalities. They used Japanese prefectural panel data for the manufacturing industry from 1991 to 2000, and concluded that spatial externalities tended to decline and became insignificant after 1993. However, it is hard to consider their examination as a long-run investigation of spatial externalities as their study period was only 10 years.

There are no definitive answers to how the extent of spatial externalities behaves in the long run as yet. In this paper, panel data for Chinese provinces from 1979 to 2003, and for Japanese prefectures from 1955 to 1998, were used to estimate the production function with the spatial externalities of China and Japan, respectively, and to investigate the long-run behavior of the spatial externalities of each country. Section 2 of this paper explains the production function including the spatial externalities. Section 3 discusses the Bayesian estimation method, and Section 4 reports the empirical results.

2. Model

Let us consider a regional economy that produces output using capital and labor input, assuming that its production technology is given by the following Cobb-Douglas form:

$$Y_{it} = A_{it} K_{it}^{\alpha_t} L_{it}^{1-\alpha_t} \exp\{\varepsilon_{it}\} \tag{1}$$

where i and t denote a region and time, Y_{it} is output, K_{it} is capital input, L_{it} is labor input, A_{it} is the level of technology, α_t is a parameter, and ε_{it} is an error term. To introduce spatial externalities into the production function, we assume the existence of externalities related to the technology level A_{it}, specifying it as follows:

$$A_{it} = \gamma_t \, \delta_t^{d_i} \prod_{j=1}^{N} \left(\frac{Y_{jt}}{L_{jt}} \right)^{\rho_t w_{ij}} \qquad (w_{ij} = 0, \text{ if } i = j)\,. \tag{2}$$

The term $\left(\frac{Y_{jt}}{L_{jt}}\right)^{\rho_t w_{ij}}$ means that increasing the labor productivity of j's region ($j \neq i$) by 1% brings about an improvement of A_{it} by $\rho_t w_{ij}$%, and this term thus indicates the spatial externalities or spatial spillover effects between region i and its neighbors. The magnitude of the spatial externalities is represented by ρ_t. The w_{ij} indicates the neighbors of region i, specifying $w_{ij} = c_{ij}/\sum_j^N c_{ij}$, where $c_{ij} = 1$ if i and j are neighbors, and $c_{ij} = 0$ otherwise. The w_{ij} refers to the standardized spatial weight ($0 \leq w_{ij} \leq 1$).

The remaining γ_t and δ_t are parameters, and d_i is China's coastal-inland dummy variable, such that $d_i = 1$ if $i \in$ coastal region, $d_i = 0$ otherwise.[1)] Consequently, the coastal and inland technology levels are distinguished such that

$$A_{it} = \begin{cases} \gamma_t \, \delta_t \prod_{j=1}^{N} \left(\frac{Y_{jt}}{L_{jt}}\right)^{\rho_t w_{ij}} & i \in \text{coastal region} \\ \gamma_t \, \prod_{j=1}^{N} \left(\frac{Y_{jt}}{L_{jt}}\right)^{\rho_t w_{ij}} & i \in \text{inland region} \end{cases}$$

In the case of Japan, we suppose that $d_i = 0$ for all i.

Substituting Equation (2) for (1), dividing by L_{it}, and taking logarithms yields the following estimable equation:

$$y_{it} = \rho_t \sum_{j=1}^{N} w_{ij} \, y_{jt} + \mathbf{x}_{it} \boldsymbol{\beta}_t + \varepsilon_{it} \tag{3}$$

where $y_{it} = \log(Y_{it}/L_{it})$, $\mathbf{x}_{it} = [1,\ d_i,\ \log(K_{it}/L_{it})]$, and $\boldsymbol{\beta}_t = [\log\gamma_t,\ \log\delta_t,\ \alpha_t]'$. Equation (3), which is called a spatial lag model in the literature (Anselin 1988; 2001), is estimated using the data of China and Japan, respectively.

In the vector and matrix notation, Equation (3) can be written as

$$\mathbf{y}_t = \rho_t \mathbf{W} \mathbf{y}_t + \mathbf{X}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t \tag{4}$$

$$\mathbf{y} = (\mathbf{D}_\rho \otimes \mathbf{W})\,\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{5}$$

where $\mathbf{y}_t = [y_{1t}, y_{2t}, \ldots, y_{Nt}]'$, $\mathbf{X}_t = [\mathbf{x}_{1t}, \mathbf{x}_{2t}, \ldots, \mathbf{x}_{Nt}]'$, $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_T]'$, and

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 & & & \\ & \mathbf{X}_2 & & \\ & & \ddots & \\ & & & \mathbf{X}_T \end{bmatrix}, \quad \mathbf{D}_\rho = \begin{bmatrix} \rho_1 & & & \\ & \rho_2 & & \\ & & \ddots & \\ & & & \rho_T \end{bmatrix}, \quad \mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \ldots & w_{1N} \\ w_{21} & w_{22} & \ldots & w_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N1} & w_{N2} & \ldots & w_{NN} \end{bmatrix}$$

in which $\mathbf{W}$ is referred to as the (row-standardized) spatial weight matrix. The reduced form of Equation (4) is given by

$$\mathbf{y}_t = (\mathbf{I}_N - \rho_t \mathbf{W})^{-1} \mathbf{X}_t \boldsymbol{\beta}_t + (\mathbf{I}_N - \rho_t \mathbf{W})^{-1} \boldsymbol{\varepsilon}_t \tag{6}$$

[1)] The coastal regions are defined as the following 12 regions: Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong, Guangxi, and Hainan. The inland regions are defined as the following 18 regions: Shanxi, Inner Mongolia, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.

where $(\mathbf{I}_N - \rho_t \mathbf{W})^{-1}$, like a Leontief inverse, can be expanded into an infinite series $(\mathbf{I}_N + \rho_t \mathbf{W} + \rho_t^2 \mathbf{W}^2 + \cdots)$, which is known in the literature as the spatial multiplier (Anselin 2003). Equation (6) implies that labor productivity in region i is affected, not only by the technology level and the capital-labor ratio in i, but also by those in all the other regions through the inverse term.

To measure the contribution of the spatial externalities (or the spatial multiplier effects) on the total production in a country, we define $Y_t = \sum_i^N Y_{it}$ as the observed total output of a country, and $\tilde{Y}_t$ as the total output without the spatial externalities:

$$\begin{aligned}
\tilde{Y}_t &= \sum\nolimits_i^N \tilde{Y}_{it} \\
&= \sum\nolimits_i^N \gamma_t \delta_t^{d_i} K_{it}^{\alpha_t} L_{it}^{1-\alpha_t} \exp\{\varepsilon_{it}\} \\
&= \sum\nolimits_i^N \exp\left\{ \log L_{it} + y_{it} - \rho_t \sum\nolimits_{j=1}^N w_{ij} y_{jt} \right\}
\end{aligned}$$

and also define the difference between Y_t and $\tilde{Y}_t$ as follows:

$$GAP_t = \frac{Y_t - \tilde{Y}_t}{Y_t} = \frac{Y_t/L_t - \tilde{Y}_t/L_t}{Y_t/L_t} \tag{7}$$

where $L_t = \sum_i^N L_{it}$ is the total labor input. The GAP_t indicates the magnitude of the spatial externalities in total domestic production. In this way, by estimating Equations (3) and (7), and describing the behavior of ρ_t and GAP_t, it is possible to investigate the long-run behavior of spatial externalities.

3. Bayesian Estimation

This section describes the Bayesian method of estimating Equations (3) and (7). Bayesian methodology requires the *posterior density* to make an inference regarding the unknown parameters in a model. The posterior is proportional to the *likelihood function* times the *prior density*, such as $\pi(\boldsymbol{\theta} \mid \mathbf{y}) \propto f(\mathbf{y} \mid \boldsymbol{\theta}) \times \pi(\boldsymbol{\theta})$, where $\mathbf{y}$ is the data observed, $\boldsymbol{\theta}$ is the unknown parameters, $\pi(\boldsymbol{\theta} \mid \mathbf{y})$ is the posterior, and $f(\mathbf{y} \mid \boldsymbol{\theta})$ is the likelihood. The following subsections explain the likelihood and the prior for our model, and show the computational scheme for estimating the posterior.

3.1 Likelihood Function

Let us assume that $\boldsymbol{\varepsilon}$ in Equation (5) has a multivariate Normal distribution, with $\mathrm{E}(\boldsymbol{\varepsilon}) = \mathbf{0}$ and $\mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \Omega_{NT}$. Then, the likelihood for our model can be expressed by

$$\begin{aligned}
f(\mathbf{y} \mid \boldsymbol{\beta}, \Omega_{NT}, \mathbf{D}_\rho) &= (2\pi)^{-\frac{NT}{2}} \mid \Omega_{NT} \mid^{-\frac{1}{2}} \prod\nolimits_{t=1}^T \mid \mathbf{I}_N - \rho_t \mathbf{W} \mid \\
&\times \exp\left\{ -\frac{1}{2} \left[\mathbf{y} - \left(\mathbf{D}_\rho \otimes \mathbf{W} \right) \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \right]' \Omega_{NT}^{-1} \left[\mathbf{y} - \left(\mathbf{D}_\rho \otimes \mathbf{W} \right) \mathbf{y} - \mathbf{X}\boldsymbol{\beta} \right] \right\}.
\end{aligned} \tag{8}$$

Since it is not feasible to estimate the $(NT \times NT)$ matrix Ω_{NT} with no restrictions, we specify the covariance matrix as follows. Suppose that $\boldsymbol{\varepsilon}_t$ follows AR(1) process

$$\boldsymbol{\varepsilon}_t = \psi \boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \Sigma_N)$$

where $|\psi| < 1$, $N()$ denotes a Normal distribution, and Σ_N is a $(N \times N)$ diagonal matrix with heteroskedasticity $(\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)$. As a result, the covariance matrix can be specified by

$$\begin{aligned} \mathrm{E}(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') &= (1-\psi^2)^{-1}\Psi_T \otimes \Sigma_N \\ &= \Omega_T \otimes \Sigma_N \\ \Psi_T &= \begin{bmatrix} 1 & \psi & \psi^2 & \ldots & \psi^{T-1} \\ \psi & 1 & \psi & \ldots & \psi^{T-2} \\ \psi^2 & \psi & 1 & \ldots & \psi^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi^{T-1} & \psi^{T-2} & \psi^{T-3} & \ldots & 1 \end{bmatrix} \end{aligned} \tag{9}$$

where $\Omega_T = (1-\psi^2)^{-1}\Psi_T$. Substituting Equation (9) into Equation (8) yields the following likelihood function:

$$\begin{aligned} f(\mathbf{y} \mid \boldsymbol{\beta}, \Sigma_N, \mathbf{D}_\rho, \psi) &= (2\pi)^{-\frac{NT}{2}} \mid \Omega_T \mid^{-\frac{N}{2}} \mid \Sigma_N \mid^{-\frac{T}{2}} \prod_{t=1}^{T} \mid \mathbf{I}_N - \rho_t \mathbf{W} \mid \\ &\times \exp\left\{-\frac{1}{2}\left[\mathbf{y} - \left(\mathbf{D}_\rho \otimes \mathbf{W}\right)\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right]' \left[\Omega_T^{-1} \otimes \Sigma_N^{-1}\right] \left[\mathbf{y} - \left(\mathbf{D}_\rho \otimes \mathbf{W}\right)\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\right]\right\}. \end{aligned} \tag{10}$$

As an alternative representation, and applying Prais-Winsten transformation to Ω_T^{-1}, Equation (10) can be rewritten as

$$\begin{aligned} f(\mathbf{y} \mid \boldsymbol{\beta}, \Sigma_N, \mathbf{D}_\rho, \psi) &= (2\pi)^{-\frac{NT}{2}} (1-\psi^2)^{\frac{N}{2}} \mid \Sigma_N \mid^{-\frac{T}{2}} \prod_{t=1}^{T} \mid \mathbf{I}_N - \rho_t \mathbf{W} \mid \\ &\times \exp\left\{-\frac{1}{2}\left[\boldsymbol{\varepsilon}_1'(1-\psi^2)\Sigma_N^{-1}\boldsymbol{\varepsilon}_1 + \sum_{t=2}^{T}(\boldsymbol{\varepsilon}_t - \psi\boldsymbol{\varepsilon}_{t-1})'\Sigma_N^{-1}(\boldsymbol{\varepsilon}_t - \psi\boldsymbol{\varepsilon}_{t-1})\right]\right\} \end{aligned} \tag{11}$$

where $\boldsymbol{\varepsilon}_t = \mathbf{y}_t - \rho_t \mathbf{W}\mathbf{y}_t - \mathbf{X}_t \boldsymbol{\beta}_t$ for $t = 1, 2, \ldots, T$.

3.2 Prior Density Function

Let us assume that the prior used in this paper takes the following form

$$\pi(\boldsymbol{\beta}, \Sigma_N, \mathbf{D}_\rho, \psi) = \pi(\boldsymbol{\beta}) \left\{\prod_{i=1}^{N} \pi(\sigma_i^2)\right\} \left\{\prod_{t=1}^{T} \pi(\rho_t)\right\} \pi(\psi) \tag{12}$$

where

$\pi(\boldsymbol{\beta})$: $\boldsymbol{\beta}_1 \sim N(\mathbf{b}_0, \Sigma_0)$, $\boldsymbol{\beta}_{t+1} = \boldsymbol{\beta}_t + \mathbf{u}_t$, $\mathbf{u}_t \sim N(\mathbf{0}, \Sigma_\beta)$
$\pi(\sigma_i^2)$: $\sigma_i^2 \sim IG(\nu_{0i}/2, \omega_{0i}/2)$ $(i = 1, 2, \ldots, N)$
$\pi(\rho_t)$: $\rho_t \sim U(\lambda_{\min}^{-1}, \lambda_{\max}^{-1})$ $(t = 1, 2, \ldots, T)$
$\pi(\psi)$: $\psi \sim TN_{(|\psi|<1)}(q_0, \sigma_{\psi 0}^2)$.

$IG()$ and $U()$ denote a distribution of inverse Gamma and Uniform, respectively. $TN_{(|\psi|<1)}$ denotes a Normal distribution, truncated on the interval $(-1 < \psi < 1)$. The hierarchical prior is introduced into $\boldsymbol{\beta}$ so that the behavior of $\boldsymbol{\beta}_t$ follows a random walk process, implying that $\boldsymbol{\beta}_t$ has a stochastic time trend. Since Σ_β is treated as an unknown parameter and requires its own prior, we assume the prior of Σ_β as $\Sigma_\beta \sim IW(\nu_{\beta 0}, \Sigma_{\beta 0})$, where $IW()$ denotes the inverse Wishart distribution.

The prior parameters are $\mathbf{b}_0$, Σ_0, ν_{0i}, ω_{0i}, $\lambda_{\min}$, $\lambda_{\max}$, q_0, $\sigma_{\psi 0}^2$, $\nu_{\beta 0}$, and $\Sigma_{\beta 0}$. The $\lambda_{\min}$ and $\lambda_{\max}$ indicate the smallest and largest eigen value of the $\mathbf{W}$, and we put a limit on the parameter space of ρ_t, such as $\lambda_{\min}^{-1} < \rho_t < \lambda_{\max}^{-1}$, which is a condition that guarantees $|\mathbf{I}_N - \rho_t \mathbf{W}| > 0$. The other prior parameters are assumed as follows:

$$\begin{aligned}
&\mathbf{b}_0 = \mathbf{0}\,, \quad \Sigma_0 = 100 * \mathbf{I}_k \\
&\nu_{0i} = 3\,, \quad \omega_{0i} = 0.01 \quad (i = 1, 2, \ldots, N) \\
&q_0 = 0.8\,, \quad \sigma_{\psi 0}^2 = (q_0/2)^2 \\
&\nu_{\beta 0} = 3\,, \quad \Sigma_{\beta 0} = 100 * \mathbf{I}_k
\end{aligned}$$

where k denotes the number of rows of $\boldsymbol{\beta}$. The $\sigma_{\psi 0}^2$ is chosen so that ψ lies in the positive area with 95.4% probability, because we expect that ε_t has a relatively high positive serial correlation.

3.3 Posterior Density Function

Having clarified the likelihood and the prior for our model, we now explain the posterior inference procedure. As is well known in the econometric literature, particularly Bayesian econometrics, the posterior inference can be carried out by the Markov Chain Monte Carlo (MCMC) method. This method allows us to generate samples from the joint posterior $\pi(\boldsymbol{\beta}, \mathbf{D}_\rho, \Sigma_N, \psi, \Sigma_\beta \mid \mathbf{y})$ and the marginal posterior of each parameter. By using the samples generated by MCMC, it can make a statistical inference about our posterior density.

The MCMC method requires us to draw samples from the *full* conditional posterior for all of the parameters, such as $\pi(\boldsymbol{\beta} \mid \mathbf{D}_\rho, \Sigma_N, \psi, \Sigma_\beta, \mathbf{y})$, $\pi(\rho_t \mid \boldsymbol{\beta}, \mathbf{D}_{-\rho_t}, \Sigma_N, \psi, \Sigma_\beta, \mathbf{y})$, $\pi(\Sigma_N \mid \boldsymbol{\beta}, \mathbf{D}_\rho, \psi, \Sigma_\beta, \mathbf{y})$, $\pi(\psi \mid \boldsymbol{\beta}, \mathbf{D}_\rho, \Sigma_N, \Sigma_\beta, \mathbf{y})$, and $\pi(\Sigma_\beta \mid \boldsymbol{\beta}, \mathbf{D}_\rho, \Sigma_N, \psi, \mathbf{y})$, where $\mathbf{D}_{-\rho_t}$ indicates the set of parameters $\rho_1, \rho_2, \ldots, \rho_T$ except for ρ_t. The method of generating samples from these full conditional distributions is discussed in Appendix B. For the estimation of GAP_t $(t = 1, 2, \ldots, T)$, we calculate the posterior mean of $\tilde{Y}_t$, using the MCMC draws such that

$$\hat{\tilde{Y}}_t = \frac{1}{R - R_0} \sum_{r=R_0+1}^{R-R_0} \left[\sum_i^N \exp\left\{ \log L_{it} + y_{it} - \rho_t^{(r)} \sum_{j=1}^N w_{ij} y_{jt} \right\} \right],$$

where R is the number of MCMC replications, R_0 is the length of burn-in period, and $\rho_t^{(r)}$ is the sample from the marginal posterior distribution of ρ_t. As mentioned in Appendix B, we set $R = 500000$ and $R_0 = 50000$, and then 450000 replications are used to calculate $\hat{\tilde{Y}}_t$, $t = 1, 2, \ldots, T$. Hence, by replacing $\tilde{Y}_t$ with $\hat{\tilde{Y}}_t$ in Equation (7), we obtain the estimate of GAP_t.

4. Estimation Results

This paper used the panel data for 30 Chinese provinces (Chongqing is included in Sichuan) from 1979 to 2003, and 46 Japanese prefectures (all except Okinawa) from fiscal years 1955 to 1998. The data description and source are reported in Appendix A. The estimation results are shown in Figures 1–4.[2)]

4.1 China

As Figure 1 demonstrates, the magnitude of spatial externalities ρ_t in China was 0.060 in 1979 and 0.183 in 2003. The ρ_t indicated a tendency to increase and statistical significant at 95% credible interval, in the period 1994–2003. However, it was insignificant before 1993. This

[2)] All computations were implemented with *Ox* version 4.04 (Doornik 2006).

indicates that the spatial externalities have appeared since 1993, and have contributed to the growth of China's economy since then.

The posterior mean of α_t, which indicates the capital elasticity, was 0.336 in 1979 and 0.468 in 2003. The capital elasticity declined between 1990 and 1994, but it showed a tendency to increase throughout the sample period, and the mean of the growth rate between 1979 and 2003 was 1.33%.The posterior mean of $\log \gamma_t$ was 0.046 in 1979 and 0.381 in 2003. The one of $\log (\gamma_t \delta_t)$ was 0.368 in 1979 and 0.770 in 2003. The γ_t and $\gamma_t \delta_t$ represent the exogenous technology level of China's inland and coastal regions, respectively. These results indicate that the coastal technology level is higher than the inland level over the sample period. In addition, the mean of the growth rate of γ_t is 1.70% (1979–90), 2.72% (1990–95), and -0.37% (1995–2003). On the other hand, that of $\gamma_t \delta_t$ is 1.65% (1979–90), 4.52% (1990–95), and -0.75% (1995-2003). The exogenous technology growth rate from 1990 to 1995 is higher in the coastal region than in the inland region.

Figure 2 shows the posterior mean of GAP_t, $\tilde{Y}_t/L_t$, and ρ_t for China. China's GAP_t was steadily increasing after 1992, the year in which Deng Xiaoping undertook his *southern tour* of China. The value of the estimated GAP_t in 1992 was 0.101, and it reached 0.355 in 2003. These results indicate that spatial externalities (or spatial multiplier effects) existed in the Chinese economy in the 1990s, and significantly contributed to China's rapid economic growth then.

4.2 Japan

Figure 3 shows Japan's estimation results. The posterior mean of ρ_t was significant from 1960 to 1974 and from 1985 to 1991, and remained insignificant during the other periods. The arithmetic mean of the estimated ρ_t over the sample period was 0.105, and its minimum value was 0.014 in 1956, and the maximum value was 0.200 in 1969. While China's ρ_t showed a rising tendency in the 1990s, Japan's ρ_t showed fluctuating values, and it was not constant over the period studied.

The posterior mean of α_t in Japan was 0.554 in 1955 and 0.568 in 1998, and its arithmetic mean over the sample period was 0.560. In comparison with that in China, capital elasticity in Japan was higher and more stable throughout the sample period. The posterior mean of $\log \gamma_t$ was -0.172 in 1955 and 0.080 in 1998. Japan's γ_t tended to increase from 1955 to 1975 (the mean of the growth rate was 1.79% in this period), and after 1975 it decreased slightly, or remained almost constant. The mean of growth rate of γ_t over the sample period in Japan was 0.56%, i.e., lower than in China.

Figure 4 displays Japan's posterior mean of GAP_t, $\tilde{Y}_t/L_t$, and ρ_t. The GAP_t and ρ_t showed an increasing phase and a decreasing phase over the period studied. The value of the estimated GAP_t was 0.229 in 1972, 0.049 in 1980, and 0.253 in 1988. The two phases are probably related to the Japanese business cycle, because GAP_t showed a high value in the period of the economic boom between 1965 and 1973 (*Izanagi boom*) and between 1986 and 1991 (*Heisei boom*), but decreased in the economic depression, due to the two *Oil crises*, in 1973 and 1979, and to the collapse of Japan's economic bubble in 1991. Taking into account the behavior of ρ_t and GAP_t for both China and Japan, it may be assumed that spatial externalities correlate with business conditions.

5. Conclusions

This paper used regional panel data for Chinese provinces from 1979 to 2003, and Japanese prefectures from 1955 to 1998, to estimate the spatial externalities (or spatial multiplier effects), using a spatial lag model and Bayesian methodology, and analyzed the long-run behavior of spatial externalities in China and Japan. According to the estimation results for China, spatial externalities significantly increased domestic production from 1994 onwards, and tended to increase until 2003. Before 1993, however, spatial externalities were insignificant.

Japan's empirical results also show that spatial externalities contributed significantly to increasing domestic production. Furthermore, the magnitude of the effects was not constant over time, but included two phases, in which they exhibited high and low values, respectively. It seems that the movement of spatial externalities is correlated with Japan's business conditions, in such a way that the externalities have a high value in an economic boom, and a low value in an economic depression. These findings lead us to presume that spatial externalities correlate mainly with business conditions.

References

Anselin L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.

Anselin L. 2001. "Spatial Econometrics." in *A Companion to Theoretical Econometrics*, Baltagi B. (eds). Malden Mass: Blackwell Publishers; 310–330.

Anselin L. 2003. "Spatial Externalities, Spatial Multipliers, and Spatial Econometrics." *International Regional Science Review* 26: 153–166.

Doi T. 2002. *Chiiki kara mita Nihon Keizai to Zaiseiseisaku* (Japanese Economy and Fiscal Policy Seen from Regions). Tokyo: The Mitsubisi Economic Research Institute. In Japanese.

Doornik JA. 2006. *Ox: An Object-Oriented Matrix Programming Language*. London: Timberlake Consultants.

Durbin J, Koopman SJ. 2002. "A Simple and Efficient Simulation Smoother for State Space Time Series Analysis." *Biometrika* 89: 603–615.

Ertur C, Koch W. 2007. "Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence." *Journal of Applied Econometrics* 22: 1033–1062.

Fingleton B, López-Bazo E. 2006. "Empirical Growth Models with Spatial Effects." *Papers in Regional Science* 85: 177–198.

Hashiguchi Y, Chen K. 2006. "Chuugoku no Shoubetsu Shihon Stock no Suikei: Ezaki-Son no Houhou to Daitaiteki Houhou (Estimating China's Provincial Capital Stocks: Ezaki and Sun's Method and an Alternative)." *Kokumin Keizai Zasshi (Journal of political economy and commercial science)* 193(6): 73–86. In Japanese.

Heston A, Summers R, Aten B. 2002. *Penn World Tables Version 6.1*. Downloadable Data Set. Center for International Comparisons of Production, Income and Prices (CIC), University of Pennsylvania.

Kakamu K, Polasek W, Wago H. 2007. "Spatial Agglomeration and Spill-over Analysis for Japanese Prefectures During 1991–2000." *MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand*, 958–964.

Kato H, Chen K. 2002. *China*, East Asian Long-Term Econometric Data Ser. 12. Tokyo Iwanami Shoten. In Japanese.

Koopman SJ, Shephard N, Doornik JA. 1999. "Statistical Algorithms for Models in State Space Using SsfPack 2.2." *Econometrics Journal* 2: 107–160.

Olejnik A. 2008. "Using the Spatial Autoregressively Distributed Lag Model in Assessing the Regional Convergence of Per-capita Income in the EU25." *Papers in Regional Science* 87: 371–384.

Pfaffermayr M. 2009. "Conditional β- and σ-Convergence in Space: A Maximum Likelihood Approach." *Regional Science and Urban Economics* 39: 63–78.

Váya E, López-Bazo E, Moreno R, Suriñach J. 2004. "Growth and Externalities Across Economies: An Empirical Analysis Using Spatial Econometrics." in *Advances in Spatial Econometrics: Methodology, Tools and Applications*, Anselin L, Florax RJGM, Rey SJ (eds). Berlin: Springer; 433–455.

Appendix A. Data Description and Source

China's data set, of Y_{it}, K_{it}, and L_{it}, is constructed as follows: Y_{it} is a provincial gross value added (Unit: million yuan) in 1990 prices, obtained by

$$Y_{it} = NY_{i,1990} \times GDPI_{it} \quad (i = 1, 2, \ldots, 30\,;\ t = 1979, 1980, \ldots, 2003)$$

where $NY_{i,1990}$ is region i's nominal gross value added in 1990, and $GDPI_{it}$ is a real GDP index at constant prices of 1990, normalized such as $GDPI_{i,1990} = 1$. $NY_{i,1990}$ is taken from *China Statistical Year Book*. $GDPI_{it}$ from 1978 to 1998 is available in Kato and Chen (2002) and the remaining data, from 1999 to 2003, is obtained from *China Statistical Year Book*. K_{it} is a provincial capital stock at 1990 prices (Unit: million yuan), obtained from Hashiguchi and Chen (2006). L_{it} is the number of provincial employed persons (Unit: 1000 persons), calculated by $L_{it} = 0.5 \times (Lye_{it} - Lye_{i,t-1})$, where Lye_{it} is the number of persons employed at the end of the year, taken from Kato and Chen (2002) for 1978 to 1998, and from the *China Statistical Year Book* for 1999 to 2003.

Japan's data set is constructed as follows. Y_{it} is the gross prefectural products at 1990 price, obtained from the *Report on the Prefectural Accounts from 1955 to 1974*, for 1955 to 1974, and from the *Annual Report on the Prefectural Accounts* for 1975 to 1998. K_{it} consists of the sum of the social and private capital stock at 1990 prices (Unit: million yen) [both figures from Doi (2002)]. L_{it} is the number of employed persons, taken from Doi (2002) for 1955 to 1974, and from the *Annual Report on the Prefectural Accounts* for 1975 to 1998 (Unit: persons).

For the specification of the spatial weight matrix $\mathbf{W}$, we used the notion of binary contiguity (Anselin 1988, pp. 18–19), assuming that regions i and j are regarded as neighbors ($c_{ij} = 1$) if they have a common border.[3)]

Appendix B. Full Conditional Posterior Density and MCMC Algorithm

The Appendix shows how to generate samples from the full conditional posterior distribution, and the MCMC algorithm.

[3)] Since Japan consists of four main islands (Hokkaido, Honshu, Shikoku, and Kyushu), these islands do not border on each other. However, as Kakamu et al. (2007) mentioned, they are connected by a bridge, tunnel, or railway. We assume that Hokkaido neighbors on Aomori (in Honshu), Hyogo (in Honshu) neighbors on Tokushima (in Shikoku), Okayama (in Honshu) neighbors on Kagawa (in Shikoku), Hiroshima (in Honshu) neighbors on Ehime (in Shikoku), and Yamaguchi (in Honshu) neighbors on Fukuoka (in Kyushu). Hainan, which is an island of China, is assumed to neighbor on Guangdong (on China's main land).

B1. Full Conditional Posterior of $\boldsymbol{\beta}$

As mentioned in section 3.2, we have assumed that the behavior of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_T$ follows a random walk process: that is, a stochastic time trend. Let us now regard $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_T$ as state variables, and exploit a state-space representation to efficiently draw $\boldsymbol{\beta}$ from the full conditional posterior. To derive the state-space form, we modify Equation (4), such that

$$\bar{\mathbf{y}}_t = \mathbf{X}_t \boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t \tag{13}$$

where $\bar{\mathbf{y}}_t = (\mathbf{I}_N - \rho_t \mathbf{W}) \mathbf{y}_t$, which is given under the full conditional distribution. By applying Prais-Winsten transformation, Equation (13) is reformulated as follows:

$$\bar{\mathbf{y}}_t^* = \begin{bmatrix} \mathbf{X}_t^* & \mathbf{MX}_t^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_t \\ \boldsymbol{\beta}_{t-1} \end{bmatrix} + \boldsymbol{\eta}_t \quad (t = 1, 2, \ldots T) \tag{14}$$

where

$$\bar{\mathbf{y}}_t^* = \begin{cases} \sqrt{1-\psi^2}\, \bar{\mathbf{y}}_1 & (t = 1) \\ \bar{\mathbf{y}}_t - \psi \bar{\mathbf{y}}_{t-1} & (t = 2, 3, \ldots, T) \end{cases}$$

$$\mathbf{X}_t^* = \begin{cases} \sqrt{1-\psi^2}\, \mathbf{X}_1 & (t = 1) \\ \mathbf{X}_t & (t = 2, 3, \ldots, T) \end{cases}$$

$$\mathbf{MX}_t^* = \begin{cases} \mathbf{0} & (t = 1) \\ -\psi \mathbf{X}_{t-1} & (t = 2, 3, \ldots, T)\,. \end{cases}$$

Then, the linear Gaussian state-space representation is given by

$$\begin{bmatrix} \boldsymbol{\beta}_{t+1} \\ \boldsymbol{\beta}_t \\ \bar{\mathbf{y}}_t^* \end{bmatrix} = \begin{bmatrix} \mathbf{I}_K & \mathbf{0} \\ \mathbf{I}_K & \mathbf{0} \\ \mathbf{X}_t^* & \mathbf{MX}_t^* \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_t \\ \boldsymbol{\beta}_{t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_t \\ \mathbf{0} \\ \boldsymbol{\eta}_t \end{bmatrix} \qquad \begin{bmatrix} \mathbf{u}_t \\ \mathbf{0} \\ \boldsymbol{\eta}_t \end{bmatrix} \sim N(\mathbf{0},\, \Omega_\beta) \qquad (t = 1, 2, \ldots, T) \tag{15}$$

$$\begin{bmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_0 \end{bmatrix} \sim N\left(\begin{bmatrix} \mathbf{b}_0 \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \Sigma_0 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \right) \qquad \Omega_\beta = \begin{bmatrix} \Sigma_\beta & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Sigma_N \end{bmatrix}.$$

Having formulated the state-space form, it is possible to exploit the simulation smoother, which is known in the literature of time series analysis (Durbin and Koopman 2001), to draw the sample from the full conditional posterior of $\boldsymbol{\beta}$. This paper used the simulation smoother developed by Durbin and Koopman (2002), using the following procedure:

Algorithm of Simulation Smoother for $\boldsymbol{\beta}$

(i) For $t = 1, 2, \ldots, T$, draw random variables $\mathbf{u}_t$ and $\boldsymbol{\eta}_t$ from $N(\mathbf{0}, \Omega_\beta)$, and use them to draw $\boldsymbol{\beta}_t$ and $\bar{\mathbf{y}}_t^*$ through Equation (15), where $\boldsymbol{\beta}_1$ is generated by $N(\mathbf{b}_0, \Sigma_0)$. The realized random variables are written by $\boldsymbol{\beta}^+ = (\boldsymbol{\beta}_1^+, \boldsymbol{\beta}_2^+, \ldots, \boldsymbol{\beta}_T^+)$ and $\bar{\mathbf{y}}^{*+} = (\bar{\mathbf{y}}_1^{*+}, \bar{\mathbf{y}}_2^{*+}, \ldots, \bar{\mathbf{y}}_T^{*+})$.

(ii) Using the simulated $\bar{\mathbf{y}}_t^{*+}$ and the real observed $\bar{\mathbf{y}}_t^*$, calculate the smoothing estimates of $\boldsymbol{\beta}$ such as $\hat{\boldsymbol{\beta}}^+ = \mathrm{E}(\boldsymbol{\beta} \mid \Sigma_N, \Sigma_\beta, \bar{\mathbf{y}}_t^{*+})$ and $\hat{\boldsymbol{\beta}} = \mathrm{E}(\boldsymbol{\beta} \mid \Sigma_N, \Sigma_\beta, \bar{\mathbf{y}}_t^*)$.

(iii) Calculate $\tilde{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}} + \boldsymbol{\beta}^+ - \hat{\boldsymbol{\beta}}^+$.

Consequently, $\tilde{\boldsymbol{\beta}}$ follows the full conditional distribution of $\boldsymbol{\beta}$. The calculation of the smoothing estimates of $\boldsymbol{\beta}$ was made by *SsfPack 2.2* (Koopman et al. 1999), which is the package of *Ox* version 4.04 programming language (Doornik 2006).

B2. Full Conditional Posterior of ρ_t, $(t = 1, 2, \ldots, T)$

The full conditional posterior of ρ_t, $(t = 1, 2, \ldots, T)$ is given by the following form

$$\pi(\rho_t \mid \boldsymbol{\beta}, D_{-\rho_t}\Sigma_N, \psi, \Sigma_\beta, \mathbf{y}) \propto |\,\mathbf{I}_N - \rho_t \mathbf{W}\,| \exp\left\{-\frac{1}{2\hat{\sigma}^2_{\rho_t}}(\rho_t - \hat{\rho}_t)^2\right\} I(\lambda_{min}^{-1} < \rho_t < \lambda_{max}^{-1}) \tag{16}$$

$$\hat{\sigma}^2_{\rho_t} = \begin{cases} \left[(\mathbf{W}\,\mathbf{y}_1)'\Sigma_N^{-1}(\mathbf{W}\,\mathbf{y}_1)\right]^{-1} & (t = 1) \\ \left[(\mathbf{W}\,\mathbf{y}_t)'\,(1+\psi^2)\,\Sigma_N^{-1}(\mathbf{W}\,\mathbf{y}_t)\right]^{-1} & (2 \le t < T) \\ \left[(\mathbf{W}\,\mathbf{y}_T)'\Sigma_N^{-1}(\mathbf{W}\,\mathbf{y}_T)\right]^{-1} & (t = T) \end{cases}$$

$$\hat{\rho}_t = \begin{cases} \hat{\sigma}^2_{\rho_1} \cdot (\mathbf{W}\,\mathbf{y}_1)'\Sigma_N^{-1}(\mathbf{y}_1 - \mathbf{X}_1\boldsymbol{\beta}_1 - \psi\,\boldsymbol{\varepsilon}_2) & (t = 1) \\ \hat{\sigma}^2_{\rho_t} \cdot (\mathbf{W}\,\mathbf{y}_t)'\Sigma_N^{-1}\left[(1+\psi^2)(\mathbf{y}_t - \mathbf{X}_t\boldsymbol{\beta}_t) - \psi(\boldsymbol{\varepsilon}_{t-1} + \boldsymbol{\varepsilon}_{t+1})\right] & (2 \le t < T) \\ \hat{\sigma}^2_{\rho_T} \cdot (\mathbf{W}\,\mathbf{y}_T)'\Sigma_N^{-1}(\mathbf{y}_T - \mathbf{X}_T\boldsymbol{\beta}_T - \psi\,\boldsymbol{\varepsilon}_{T-1}) & (t = T). \end{cases}$$

where $I(\lambda_{min}^{-1} < \rho_t < \lambda_{max}^{-1})$ is an indicator function that is equal to 1 if ρ_t lies inside the interval between λ_{min}^{-1} and λ_{max}^{-1}, and is equal to 0 otherwise. Since the density is not standard, we use the Metropolis-Hastings (MH) algorithm to draw a sample from Equation (16). The algorithm takes the following procedure:

MH Algorithm for ρ_t, $(t = 1, 2, \ldots, T)$
Suppose that r is the number of times of MCMC sampling, and choose an arbitrary starting value $\rho_t^{(r)}$ $(r = 0)$.

(i) Draw ρ_t^*, as a candidate of $\rho_t^{(r)}$, from the candidate generating density $q(\rho_t^* \mid \rho_t^{(r-1)})$.

(ii) Calculate an acceptance probability $\alpha(\rho_t^*, \rho_t^{(r-1)})$.

(iii) Set $\rho_t^{(r)} = \rho_t^*$ with probability $\alpha(\rho_t^*, \rho_t^{(r-1)})$, and set $\rho_t^{(r)} = \rho_t^{(r-1)}$ with probability $1 - \alpha(\rho_t^*, \rho_t^{(r-1)})$.

As the candidate generating density, we exploit $TN_{(\lambda_{min}^{-1} < \rho_t < \lambda_{max}^{-1})}(\hat{\rho}_t, \hat{\sigma}_t^2)$, which denotes a Normal distribution truncated on the interval $\lambda_{min}^{-1} < \rho_t < \lambda_{max}^{-1}$, and consequently the acceptance probability results in

$$\alpha(\rho_t^*, \rho_t^{(r-1)}) = \min\left\{1, \frac{|\,\mathbf{I}_N - \rho_t^*\,\mathbf{W}\,|}{|\,\mathbf{I}_N - \rho_t^{(r-1)}\,\mathbf{W}\,|}\right\}.$$

B3. Full Conditional Posterior of ψ

The full conditional posterior of $\boldsymbol{\psi}$ is given by

$$\pi(\psi \mid \boldsymbol{\beta}, D_\rho, \Sigma_N, \Sigma_\beta, \mathbf{y}) \propto A(\psi) \times \exp\left\{-\frac{1}{2\sigma^2_{\psi 1}}\,(\psi - q_1)^2\right\} I(|\,\psi\,| < 1) \tag{17}$$

where $I(|\,\psi\,| < 1)$ is an indicator function that is equal to 1 if $|\,\psi\,| < 1$, and is equal to 0 otherwise, and

$$A(\psi) = (1-\psi^2)^{\frac{N}{2}} \exp\left\{-\frac{1}{2}\left[\boldsymbol{\varepsilon}_1'(1-\psi^2)\Sigma_N^{-1}\boldsymbol{\varepsilon}_1\right]\right\}$$

$$\sigma^2_{\psi 1} = \left(\sum\nolimits_{t=2}^{T} \boldsymbol{\varepsilon}_{t-1}'\Sigma_N^{-1}\boldsymbol{\varepsilon}_{t-1} + \sigma_{\psi 0}^{-2}\right)^{-1}$$

$$q_1 = \sigma^2_{\psi 1}\left(\sum\nolimits_{t=2}^{T} \boldsymbol{\varepsilon}_{t-1}'\Sigma_N^{-1}\boldsymbol{\varepsilon}_t + \sigma_{\psi 0}^{-2}\,q_0\right).$$

The density is also not standard in the case of ρ_t, and hence we use the MH algorithm described above. We adopt $N(q_1, \sigma^2_{\psi 1})$ as the candidate generating distribution, and then the acceptance probability takes the following form:

$$\alpha_\psi(\psi^*, \psi^{(r-1)}) = \min\left\{1, \frac{A(\psi^*)}{A(\psi^{(r-1)})}\right\}.$$

B4. Full Conditional Posterior of Σ_N and Σ_β

Lastly, the full conditional posterior of Σ_N, which is the diagonal matrix of $(\sigma_1^2, \sigma_2^2, \ldots, \sigma_N^2)$, and Σ_β, takes a form such as

$$\sigma_i^2 \mid \boldsymbol{\beta}, D_\rho, \psi, \Sigma_\beta, \mathbf{y} \sim IG\left(\frac{\nu_{1i}}{2}, \frac{\omega_{1i}}{2}\right) \qquad (i = 1, 2, \ldots, N) \tag{18}$$

$$\nu_{1i} = \nu_{0i} + T$$

$$\omega_{1i} = \left((1-\psi^2)\varepsilon_{i1}^2 + \sum\nolimits_{t=2}^{T}(\varepsilon_{it} - \psi\varepsilon_{i,t-1})^2\right) + \omega_{0i}$$

$$\Sigma_\beta \mid \boldsymbol{\beta}, D_\rho, \Sigma_N, \psi, \mathbf{y} \sim IW\left(\nu_{\beta 1},\ \Sigma_{\beta 1}\right) \tag{19}$$

$$\nu_{\beta 1} = \nu_{\beta 0} + T - 1$$

$$\Sigma_{\beta 1} = \left[\sum\nolimits_{t=1}^{T-1}(\boldsymbol{\beta}_{t+1} - \boldsymbol{\beta}_t)(\boldsymbol{\beta}_{t+1} - \boldsymbol{\beta}_t)' + \Sigma_{\beta 0}^{-1}\right]^{-1}$$

where $\varepsilon_{it} = y_{it} - \rho_t \sum_{j=1}^{N} w_{ij} y_{jt} - \mathbf{x}_{it}\boldsymbol{\beta}_t$.

B5. MCMC Algorithm

Now, we show the MCMC algorithm that obtains samples from the posterior distribution.

MCMC Algorithm

Suppose that r is the number of times of MCMC sampling.

(i) Choose the arbitrary initial value for all parameters and set up $r = 1$.

(ii) Repeat the following sampling:

Draw $\boldsymbol{\beta}^{(r)}$ from $\pi(\boldsymbol{\beta} \mid \rho_1^{(r-1)}, \rho_2^{(r-1)}, \ldots, \rho_T^{(r-1)}, \Sigma_N^{(r-1)}, \psi^{(r-1)}, \Sigma_\beta^{(r-1)}, \mathbf{y})$.

Draw $\rho_1^{(r)}$ from $\pi(\rho_1 \mid \boldsymbol{\beta}^{(r)}, \rho_2^{(r-1)}, \rho_3^{(r-1)}, \ldots, \rho_T^{(r-1)}, \Sigma_N^{(r-1)}, \psi^{(r-1)}, \Sigma_\beta^{(r-1)}, \mathbf{y})$.

Draw $\rho_2^{(r)}$ from $\pi(\rho_2 \mid \boldsymbol{\beta}^{(r)}, \rho_1^{(r)}, \rho_3^{(r-1)}, \ldots, \rho_T^{(r-1)}, \Sigma_N^{(r-1)}, \psi^{(r-1)}, \Sigma_\beta^{(r-1)}, \mathbf{y})$.

$\vdots$

Draw $\rho_T^{(r)}$ from $\pi(\rho_T \mid \boldsymbol{\beta}^{(r)}, \rho_1^{(r)}, \rho_2^{(r)}, \ldots, \rho_{T-1}^{(r)}, \Sigma_N^{(r-1)}, \psi^{(r-1)}, \Sigma_\beta^{(r-1)}, \mathbf{y})$.

Draw $\psi^{(r)}$ from $\pi(\psi \mid \boldsymbol{\beta}^{(r)}, \rho_1^{(r)}, \rho_2^{(r)}, \ldots, \rho_T^{(r)}, \Sigma_N^{(r-1)}, \Sigma_\beta^{(r)}, \mathbf{y})$.

Draw $\Sigma_N^{(r)}$ from $\pi(\Sigma_N \mid \boldsymbol{\beta}, \rho_1^{(r)}, \rho_2^{(r)}, \ldots, \rho_T^{(r)}, \psi^{(r)}, \Sigma_\beta^{(r-1)}, \mathbf{y})$.

Draw $\Sigma_\beta^{(r)}$ from $\pi(\Sigma_\beta \mid \boldsymbol{\beta}^{(r)}, \rho_1^{(r)}, \rho_2^{(r)}, \ldots, \rho_T^{(r)}, \Sigma_N^{(r)}, \psi^{(r)}, \mathbf{y})$.

If $r < R$, set $r = r + 1$ and return to (ii). Otherwise, go to (iii).

(iii) Discard the draws with the superscript $r = 1, 2, \ldots, R_0$, and save the draws with $r = R_0 + 1, R_0 + 2, \ldots, R$.

In this paper, we take $R = 500000$ and $R_0 = 50000$, and then 450000 replications are retained and exploited to implement the posterior inference.

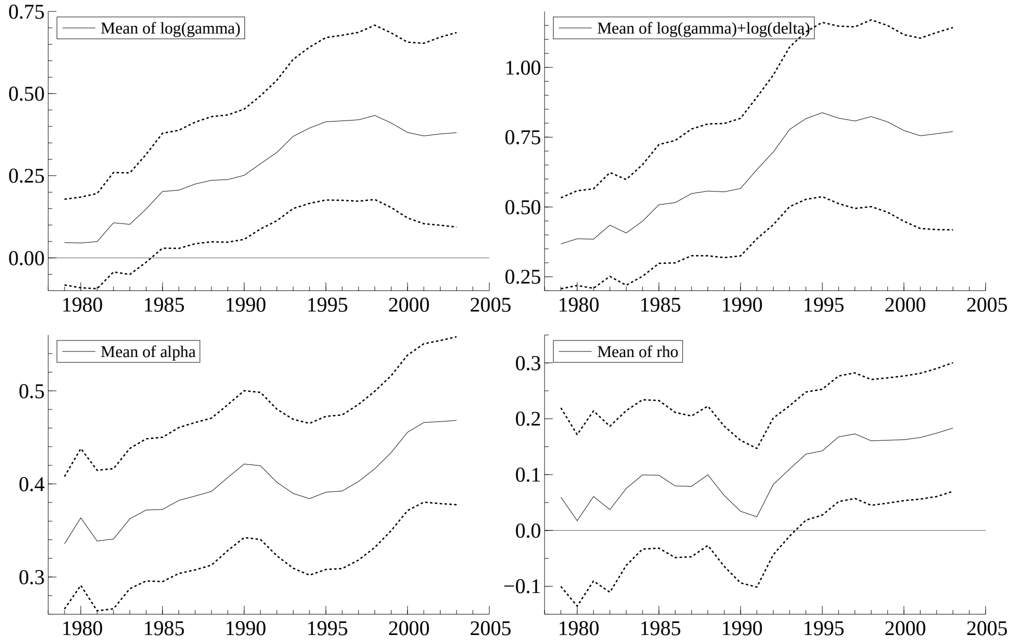

Figure 1: Posterior Mean and 95% Credible Interval (China)

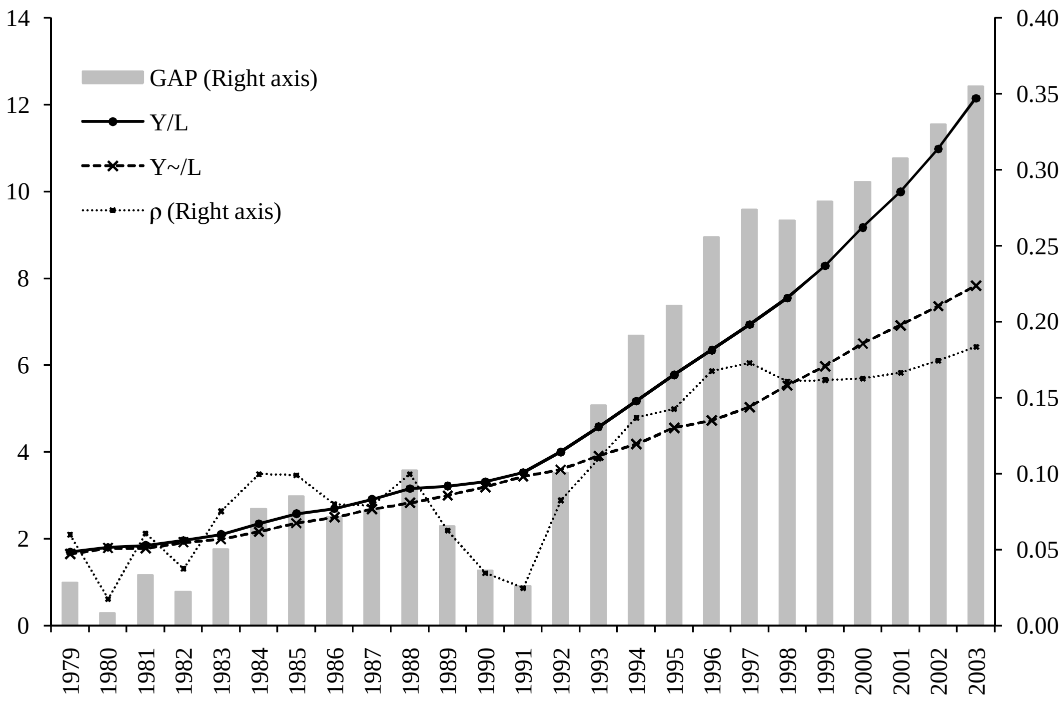

Figure 2: The Difference Between Y/L and $\tilde{Y}/L$ (China)

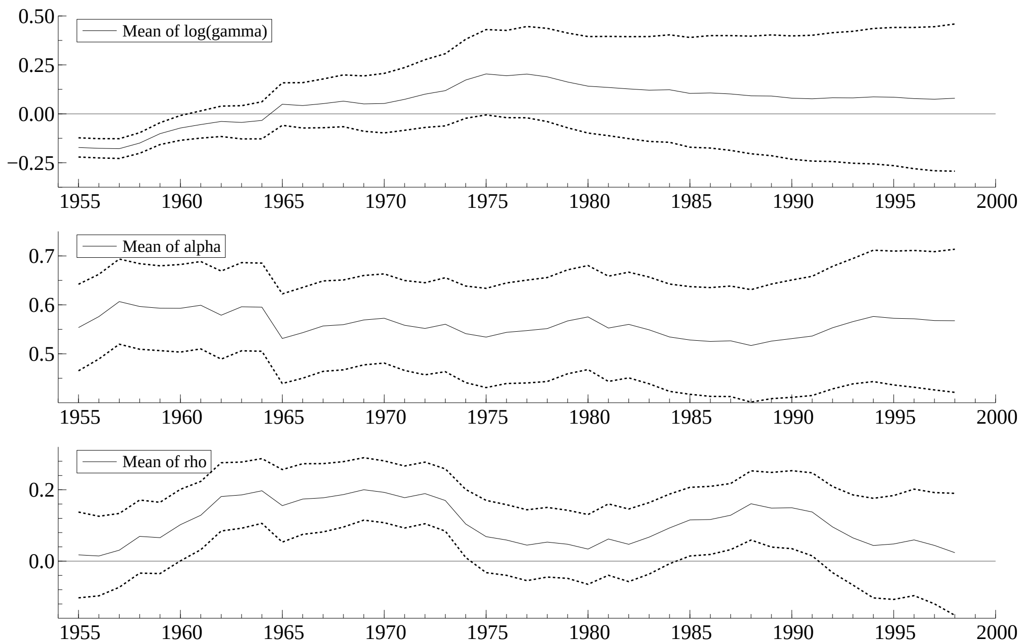

Figure 3: Posterior Mean and 95% Credible Interval (Japan)

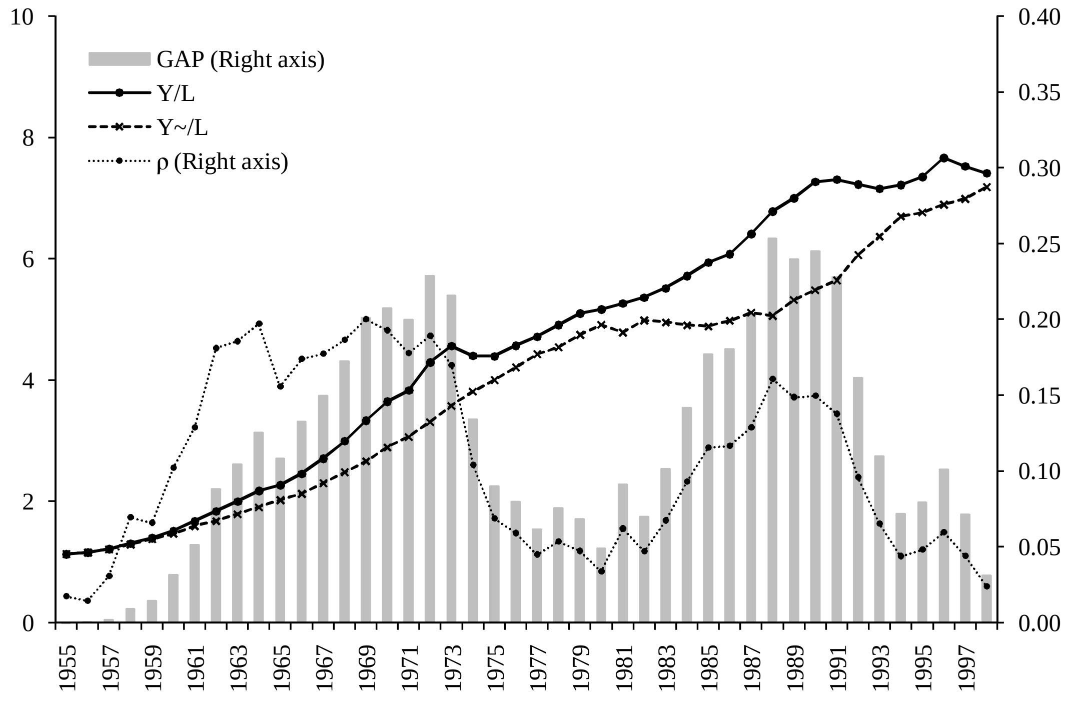

Figure 4: The Difference Between Y/L and $\tilde{Y}/L$ (Japan)