DSSR

Discussion Paper No. 81

# Prefecture-Level Spatio-temporal Analysis of Foreign Labour in Japan

Junyue Wu
Yasumasa Matsuda

April 17, 2018

Data Science and Service Research
Discussion Paper

Center for Data Science and Service Research
Graduate School of Economic and Management
Tohoku University
27-1 Kawauchi, Aobaku
Sendai 980-8576, JAPAN

# Prefecture-Level Spatio-temporal Analysis of Foreign Labour in Japan

Junyue Wu* and Yasumasa Matsuda*

April 17, 2018

**Abstract**

The number of foreign workers in Japan has been growing rapidly in recent years, and this phenomenon is drawing more attention than ever. Prefecture-level data shows uneven distribution across the country, but few research has been done to explain this fact. In this thesis, We proposed a threshold spatio-temporal model and its estimation method based on existing spatial panel model, attempting to improve the goodness of fit and making the comparison of prefectures with different economic properties possible. Then using a series of data published by Japanese government, We fitted the threshold model and traditional panel models to examine the influential factors of foreign labours in Japan at prefecture level. After that, We interpreted the result based on preliminary researches. As a result, the threshold model is proved superior to the two classic models in this study. Judging from the fitted models, the existence of spatial effect is confirmed, and GPP is identified as a major factor of foreign labour; enrolled students, criminal offenses and number of hotel guests as minor factor because they only showed influence in certain regions.

**Keywords**: foreign workers, spatial panel model, prefecture-level

## 1 Introduction

Japan has been attracting a significant number of foreign workers in recent years. In the year 2016, over one million foreigners were working in Japan, reaching historical high and more than doubling the number of 2008. More than two-thirds of the foreign workers came from Asia, and most of them worked in manufacturing, lodging, food service, wholesale and retail industry (Ministry of Health, Labour and Welfare, 2017 [1]). However, they still only made up about 1.68% of the workforce in Japan, less than most other industrialized countries (Statistic Bureau, Ministry of Internal Affairs and Communications, 2017 [2]).

Several explanations for the increase of foreign workers can be found in earlier studies, here lists a few of them. Firstly, the number of migrants around the world had continuously been increasing since 1990, and their employment rate was around 90% in Asia on average (UN-DESA and O. E. C. D, 2013 [3]). This result showed the increase in foreign labour is a global phenomenon. Secondly, the declining birth rate combined with the growing population of the senior citizens in Japan resulted in a fall in total population as well as workforce, consequently raising the demand for alternative labour force in order to maintain Japanese economy and society (Tezuka, 2005 [4]). Thirdly, in the age of global war of talent, Japanese government put forward its migration policy to attract highly skilled foreigners to work and live in Japan just like many other nations did. The goal was achieved to an

*Graduate School of Econometrics and Management, Tohoku University

extent, at the same time the practice also brought in other kinds of foreign workers (Chiavacci, 2012 [5]). Finally, there should be significantly more international workers migrates to Japan according to the result of international migration push/pull model which was based on the relative economic strength of Japan and other countries (Vogt, 2015 [6]).

Some empirical panel data studies had examined the relationship between immigrants and host countries economic and social properties. Although foreign workers do not equal to immigrants, the existence of similar links is highly plausible. First, in the case of international migration, emigration rate was positively related to destination countrys GDP in OECD members. Second, some geographic and cultural determinants like distance had significant influence on emigration rate, but others like common language or past colonial relationships did not (Mayda, 2010 [7]). Thirdly, the positive relation between immigration and international tourism had been observed in Australia, both in short run and long run (Seeteram, 2010 [8]). Fourthly, the share of foreign citizens was positively associated with criminal activities, especially crimes against property in Germany (Entorf et al., 2000 [9]). Finally, a close link was found between international students and migration flows, an increase in the number of foreign students leads to a rise in immigration (Dreher and Poutvaara, 2011 [10]).

Spatial econometrics models have not been widely used in foreign labour studies. In this study, to uncover the determinants of foreign labour in Japan at prefecture level, We performed a spatio-temporal analysis as well as dynamic panel analysis for comparison. In addition, We proposed a threshold spatio-temporal model based on existing model, attempting to improve the goodness of fit and making it possible to compare prefectures with different economic properties. After model estimation using data published by Japanese government, We interpreted the result based on preliminary researches.

# 2 Preliminary models

## 2.1 Dynamic panel model

First, We considered the dynamic panel model which is specified as follows (Hsiao, 2014 [11])

$$y_{it} = \gamma y_{i,t-1} + x_{it}\beta + \mu_i + \epsilon_{it}$$

where $y_{it}$ denotes the dependent variable for region $i$ at time $t$; $y_{i,t-1}$ is the lagged dependent variable and $\gamma$ is the corresponding parameter which is restricted to the interval $(-1, 1)$ under the assumption that the model is stable; $x_{it}$ is a $1 \times K$ vector of explanatory variables, and $\beta$ is a $K \times 1$ vector of the corresponding parameters; $\mu_i$ is the time-invariant fixed effect; $\epsilon_{it}$ is assumed to be normally distributed which follows $\epsilon_{it} \sim N(0, \sigma^2)$.

The parameters can be estimated by the within estimator. First, to concentrate out the individual effect, demean the dependent variable by

$$y_{it}^* = y_{it} - \frac{1}{T}\sum_{i=1}^{T} y_{it}$$

and do the same for $x_{it}$, $y_{i,t-1}$. Then the original equation becomes

$$y_{it}^* = \gamma y_{i,t-1}^* + x_{it}^*\beta + \epsilon_{it}$$

and it can be estimated by maximum likelihood. Under the assumption that $\epsilon_{it}$ is normally distributed, the log-likelihood function is

$$LogL = -\frac{NT}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(y_{it}^* - \gamma y_{i,t-1}^* - x_{it}^*\beta)^2$$

where N is the number of regions and T is the number of time periods.

## 2.2 Spatial model

Since prefectures are connected geographically, there could be spatial correlations among the prefecture data. To take spatial effect into account, a spatial model is required. Such model has higher explanatory power than conventional panel models when spatial effect exists. To explain what is a spatial model, spatial weight matrix must be introduced first.

### 2.2.1 Spatial model

Spatial weight matrix, or simply weight matrix, is a $N \times N$ matrix that represents the spatial relations among N regions in spatial models. Its elements $W_{ij}$ reflect the influents of region $j$ on region $i$. Self-influences are excluded so all the diagonal elements $W_{ii}$ equal to zero. Here is a simple example: let $N(i)$ be a set of regions that share borders with region $i$, then the elements of the weight matrix defined by border share cane be specified as

$$W_{ij} = \begin{cases} 1, & \text{if } j \in N(i) \\ 0, & \text{otherwise} \end{cases}$$

and diagonal elements $W_{ii}$ equal to zero. Usually, weight matrix is row-normalized by

$$W_{ij}^* = \frac{W_{ij}}{\sum_{i=1}^{N} W_{ij}}$$

### 2.2.2 Spatial lag model

Using the spatial weight matrix, a spatial lag term $\sum_{i=1}^{N} W_{ij}y_{it}$ can be constructed to represent spatial effect. Spatial lag model, a cross-section model with spatial lag is specified as (Anselin, 2013 [12])

$$y_i = \delta\sum_{j=1}^{N} W_{ij}y_i + x_i\beta + \epsilon_i$$

where $W$ is a $N \times N$ row-normalized weight matrix defined by border share, which means its element $W_{ij}$ is 1 if prefecture share border, otherwise 0; $\delta$ is the spatial lag parameter which measures the strength of the spatial effect, and $|\delta|$ should be less than 1 under the assumption that the model is not explosive.

The parameters can be estimated by maximizing the following log-likelihood function:

$$LogL = -\frac{N}{2}\log(2\pi\sigma^2) + \log|I_N - \delta W| - \frac{1}{2\sigma^2}\sum_{i=1}^{N}(y_i - \delta\sum_{j=1}^{N} W_{ij}y_j - x_i\beta)^2$$

### 2.2.3 Spatial dynamic panel data model

Since panel data analysis is required in this study, apparently the spatial lag model is not capable and need to be extended. The dynamic spatial lag panel data model with cross-sectional fixed effect is a combination of the dynamic panel model and the spatial lag model, which is specified as follows (Yu et al., 2008 [13]; Elhorst [14], 2010)

$$y_{it} = \gamma y_{i,t-1} + \delta \sum_{j=1}^{N} W_{ij} y_{it} + x_{it}\beta + \mu_i + \epsilon_{it}$$

where the variables and parameters have the same meaning as the previous models.
For estimation, the least-squares dummy variables (LSDV) estimator can be used. First, demean $y_{it}$, $x_{it}$, $y_{i,t-1}$, and $\sum_{j=1}^{N} W_{ij} y_{it}$ using the same method in the within estimator to concentrate out the individual effect. Then the equation

$$y^*_{it} = \gamma y^*_{i,t-1} + \delta \Big[ \sum_{j=1}^{N} W_{ij} y_{it} \Big]^* + x^*_{it}\beta + \epsilon_{it}$$

can be estimated by maximum likelihood similar to the previous model. The log-likelihood function for the demeaned equation is

$$LogL = -\frac{NT}{2} \log(2\pi\sigma^2) + T \log|I_N - \delta W| - \frac{1}{2\sigma^2} \sum_{i=1}^{N} \sum_{t=1}^{T} (y^*_{it} - \gamma y^*_{i,t-1} - \delta \Big[ \sum_{j=1}^{N} W_{ij} y_{jt} \Big]^* - x^*_{it}\beta)^2$$

# 3 Threshold model

## 3.1 Model specification

To compare the differences of prefectures and improve the goodness of fit, We made some modification to the spatial dynamic panel data model. The model is a spatial dynamic panel model divided into two parts by a predetermined variable $Z_{it}$ and a threshold value $A$, which is specified as follows:

$$y_{it} = \begin{cases} \gamma_1 y_{i,t-1} + \delta_1 \sum_{j=1}^{N} W_{ij} y_{it} + x_{it}\beta_1 + \mu_i + \epsilon_{it} & \epsilon_{it} \sim N(0, \sigma_1^2), \quad \text{when } Z_{it} \geq A \\ \gamma_2 y_{i,t-1} + \delta_2 \sum_{j=1}^{N} W_{ij} y_{it} + x_{it}\beta_2 + \mu_i + \epsilon_{it} & \epsilon_{it} \sim N(0, \sigma_2^2), \quad \text{when } Z_{it} < A \end{cases}$$

where $\gamma_1$, $\gamma_2$ are parameters for time lag $y_{i,t-1}$; $\delta_1$, $\delta_2$ are the spatial lag parameters; $\beta_1$, $\beta_2$ are $K \times 1$ vectors of the corresponding parameters for explanatory variables.

## 3.2 Estimation

The estimation method is based on LSDV estimator. The demeaning process is the same as the spatial dynamic panel model and the same variables $y^*_{it}$, $x^*_{it}$, $y^*_{i,t-1}$, and $\Big[ \sum_{j=1}^{N} W_{ij} y_{jt} \Big]^*$ are obtained. On the other hand, due to the existence of the threshold, the likelihood function consists two parts corresponding to the threshold model, and the Jacobian term need adjustment. Depending on whether $Z_{it}$ is time-dependent, the estimation method will be a little different.

### 3.2.1 Time-dependent $Z_i$ case

When $Z_{it}$ is time-dependent, the Jacobian term may vary for different time period. First, define a block diagonal matrix $B$ by

$$B = I_T \otimes W$$

where $\otimes$ denotes Kronecker product. Next, modify $B$ by

$$B^*_{it} = \begin{cases} \delta_1 B_{it}, & \text{when } Z_{it} \geq A \\ \delta_2 B_{it}, & \text{when } Z_{it} < A \end{cases}$$

and the Jacobian term can be expressed as $\log|I_N - B^*|$. Then the parameters can be estimated by maximizing the following likelihood function:

$$\begin{aligned} LogL = & -\frac{N_1 T}{2}\log(2\pi\sigma_1^2) - \frac{N_2 T}{2}\log(2\pi\sigma_2^2) + \log|I_N - B^*| \\ & - \frac{1}{2\sigma_1^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(y^*_{it} - \gamma_1 y^*_{i,t-1} - \delta_1\Big[\sum_{j=1}^{N} W_{ij}y_{jt}\Big]^* - x^*_{it}\beta_1)^2 \times I\{Z_{it\geq A}\} \\ & - \frac{1}{2\sigma_2^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(y^*_{it} - \gamma_2 y^*_{i,t-1} - \delta_2\Big[\sum_{j=1}^{N} W_{ij}y_{jt}\Big]^* - x^*_{it}\beta_2)^2 \times I\{Z_{it<A}\} \end{aligned}$$

where

$$I\{Z_{it\geq A}\} = \begin{cases} 1, & \text{when } Z_{it} \geq A \\ 0, & \text{when } Z_{it} < A \end{cases}$$

$$I\{Z_{it<A}\} = \begin{cases} 1, & \text{when } Z_{it} < A \\ 0, & \text{when } Z_{it} \geq A \end{cases}$$

and $N_1$ is the number of regions which have $Z_{it} \geq A$, $N_2$ is the number of regions which have $Z_{it} < A$.

### 3.2.2 Time-independent $Z_i$ case

When $Z_{it}$ is time-independent, it can be denoted as $Z_i$for any specific time period. As a result, the Jacobian term is also the same for all the time periods. Define a modified version of weight matrix $W^*$ by

$$W^*_{ij} = \begin{cases} \delta_1 W_{ij}, & \text{when } Z_i \geq A \\ \delta_2 W_{ij}, & \text{when } Z_i < A \end{cases}$$

and the Jacobian term is $T\log|I_N - W^*|$in this case. Then the log-likelihood function becomes

$$\begin{aligned} LogL = & -\frac{N_1 T}{2}\log(2\pi\sigma_1^2) - \frac{N_2 T}{2}\log(2\pi\sigma_2^2) + T\log|I_N - W^*| \\ & - \frac{1}{2\sigma_1^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(y^*_{it} - \gamma_1 y^*_{i,t-1} - \delta_1\Big[\sum_{j=1}^{N} W_{ij}y_{jt}\Big]^* - x^*_{it}\beta_1)^2 \times I\{Z_{it\geq A}\} \\ & - \frac{1}{2\sigma_2^2}\sum_{i=1}^{N}\sum_{t=1}^{T}(y^*_{it} - \gamma_2 y^*_{i,t-1} - \delta_2\Big[\sum_{j=1}^{N} W_{ij}y_{jt}\Big]^* - x^*_{it}\beta_2)^2 \times I\{Z_{it<A}\} \end{aligned}$$

where the variables and parameters have the same meaning as the previous models.

### 3.3 Threshold value and the corresponding variable

The variable can be the dependent variable, one of the independent variable, or even a variable beyond the dataset as long as it is a property of the regions in the research with explainable economic meaning. When given a variable, the optimal threshold value can be found by comparing the result of the fitted model, the one leads to the best fit should be selected. The optimal variable can also be chosen by comparing the goodness of fit given their optimal threshold value.

## 4 Data

In this study, prefecture-level panel data from 2008 to 2013 were applied to identify influential factors of the population of foreign workers in Japan. Due to a change in reporting policy for foreign workers in 2007, data before 2008 were discarded for consistency problems, and prefecture-level data after the year 2013 were not available for the majority of the explanatory variables. The dependent variable was the number of foreign workers in each prefecture (NFW). The independent variables were chosen for the following reasons.

Gross prefecture product (GPP): Gross prefecture product reflects the general economic strength of a prefecture. As shown in previous international migration study, immigration is usually related to GDP, and a positive relationship is expected (Mayda, 2010 [7]). GPP can be regarded as GDP at prefecture-level, so a similar relation between foreign workers and GPP is expected.

Consumer price index (CPI): Consumer price index shows the cost of living in a prefecture. Existing study stated that it is essential to include variables measuring the cost of living in migration regressions in order to avoid misspecification problems (Renas and Kumar, 1978 [15]), the same could be applied to foreign labour research.

Students enrolled in colleges and universities (SCU): The number of students enrolled in colleges and universities reflects the strength in education of a prefecture. In 2008, the government of Japan announced the 300,000 Foreign Students Plan aiming to increase the number of foreign students in Japan. Furthermore, they encourage them to work in Japan after graduation. At the same time, many foreign students have part-time jobs so they may also count as foreign workers. These may cause a correlation between educational strength and foreign labour, which was examined in this study.

Recognitions of criminal offenses (RCO): Recognitions of criminal offense represents overall safety level of a prefecture. Also, there is a general concern that foreign residents may have higher crime rate than local people, We will examine the existence of this relationship in this study.

Total hotel guests (THG): The number of total hotel guests shows the strength of culture and tourism industry of a prefecture. In addition, relationships between tourism and migration had been discussed in earlier study (Williams and hall, 2000 [16]), similar links could exist between tourism and foreign labour.

The data on the number of foreign workers were extracted from annual reports on foreign labour situation published by The Ministry of Health, Labour, and Welfare. Data on all of the independent variables were drawn from the regional statistic database provided by the statistics bureau. A logarithm transformation was performed on all of the data before model estimation. The Descriptive statistics for the variables can be found in Table 1.

# 5 Results

## 5.1 Empirical model estimation and comparison

In this section, We compare the goodness of fit and how well the spatial effect is handled for dynamic panel model, spatial dynamic panel model, and threshold spatial dynamic panel model. In the fitted threshold model, the selected variable was the growth rate of GPP from 2008 to 2013, and the chosen threshold value was 2.5%. The threshold was time-independent, and it separated the prefectures into two groups, as shown in Figure 1. Table 2 shows the estimation results, and Table 2 shows the result of Morans I test for residuals.

The goodness of fit was valued by Akaike Information Criterion (AIC). AIC measures the relative quality of each model given the dataset, and smaller AIC means a better fit. In Table 2, the AIC of dynamic panel model was significantly larger than the other two models, and threshold spatial dynamic panel model had notably smaller AIC than its non-threshold counterpart. Judging from these facts, the threshold spatial dynamic panel model was considered the best model of the three.

How well the spatial effect was handled was assessed by Morans I test. Morans I is an index that measures spatial autocorrelation (Cliff and Ord, 1981), its value range from -1 to 1. A positive value indicates positive spatial autocorrelation given it is significant, and vice versa. In Table 3, at 5% confidence interval, the residual of fitted dynamic panel model showed significant spatial autocorrelation in the year 2011 and 2013, spatial dynamic panel model had spatial autocorrelation in the year 2010, and the one of threshold spatial dynamic panel model showed no significant spatial autocorrelation across the board. So it could be said the threshold spatial dynamic panel model performed the best dealing with spatial effect in this dataset.

## 5.2 Spatio-temporal patterns of foreign workers

### 5.2.1 Spatial effect

The quantities and distribution of foreign workers in Japan were displayed in Figure 2 for the six years. During this period, the number of foreign workers increased steadily across the board, except for the year 2012, which might be a result of the Great East Japan Earthquake in 2011. Also, it was clear from the colored map that the distribution was quite uneven spatially. For example, prefectures near big cities like Tokyo, Osaka and Nagoya tended to have more foreign workers than other regions. This evidence indicated that the existence of the spatial effect seemed highly plausible.

To determine whether the spatial effect existed in the foreign worker data or not, We calculated its Morans I index and its significance level over the six years. The result can be found in Table 3, the indexes were positive and significant across the board and relatively stable in this period. At the same time, the spatial factor in the fitted spatial dynamic panel model was also positive and significant, further confirmed the existence of positive spatial relationship and showed that the spatial effect was relatively similar during the six-year period.

Moreover, the threshold spatial dynamic panel model provided a comparison for spatial effect between prefectures having high GPP growth rate and the others. In Table 2, the spatial factor for prefectures having high GPP growth was positive and significant, but the one for the other group was not. This result showed that in high GPP growth regions the spatial effect was positive and fairly strong; in prefectures with low GPP growth rate, no significant spatial effect was detected. Since the spatial weight matrix was defined by border share, it could be said that the number of foreign workers for prefectures with high GPP growth rate was heavily influenced by the ones for the surrounding regions. Nevertheless, similar impact was not observed in other prefectures.

### 5.2.2 Temporal effect

The existence of temporal effect was examined by the dynamic factor in the fitted models. The positive and significant dynamic factor in spatial dynamic panel model indicated positive temporal effect when all the prefectures were considered as a whole. On the other hand, when prefectures with high and low GPP growth rate were measured separately, the dynamic factors in the fitted threshold model showed that the temporal effect is only significant in low GPP growth regions, on the contrary of the result for spatial effect.

## 5.3 Influential factors of foreign workers

In this section, We examine how the selected explanatory variables, namely GPP, CPI, influence foreign labour in Japan by the results of three fitted models, mainly focus on ones with spatial effect.

All the fitted models showed that GPP has significant positive impact on the number of foreign workers. Furthermore, the corresponding parameters in the models were relatively similar, indicated that spatial spillover effects had little effect on the relationship between GPP and foreign labour, and the threshold model showed that the difference between high GPP growth prefectures and the low ones was minimum. This result was in line with the one in previous international migration studies as expected (Mayda, 2010), GPP was a major determinant of foreign workers in Japan, regardless of the economic strength of a prefecture.

In the case of CPI, although the parameters are negative, none of them were significant even at 10% level, which indicated no relationship between CPI and foreign labour was detected in any of the three models. This result was on the contrary to the one of earlier migration decision study, which stated a negative correlation between the cost of living and immigration was expected (Renas and Kumar, 1978). It seems that the cost of living was not a major concern in foreign workers decision-making process, or at least overshadowed by other factors.

For the number of students enrolled in colleges and universities, the parameter in the non-threshold spatial model was positive and significant at 5% level, revealing the positive link in general. On the other hand, the threshold model showed a difference between high GPP growth prefectures and the others. In prefectures with high GPP growth rate, a significant correlation was observed, but not in other regions. As stated in chapter 4, due to the policy change and the increase of international students, a positive relationship between educational strength and foreign labour was anticipated. In fact the prefectures with high GPP growth rate did not necessarily have top-level educational strength in Japan, which might seem weird at first glance. To address that, We came up with two theories: first, foreign student in these prefectures were more prone or encouraged to get a job in local establishments; second, the number of students could represent a prefectures strength that wont change over time other than education, and it might be attractive to some foreign workers.

The fitted spatial model did support the theory that immigrants might cause more crime than local people on a large scale, with a positive and significant parameter for criminal offenses. However, inequality between high GPP growth regions and the others was revealed by the threshold model. Significant relationship was only found in areas with low GPP growth rate. According to earlier studies, foreigner crime was usually linked with unemployment and prejudice (Entorf and Horst, 2000). So less educated unskilled foreign workers were more prone to commit crime than highly skilled ones, the difference in the proportion of various types foreign worker might be the cause of the phenomenon.

The impact of the number of total hotel guests on foreign labour was only present in prefectures with low GPP growth and barely significant, only at 10% level based on the result of threshold model. Usually the relationship between tourism and labour migration was mainly caused by visits of families and friends, tourism-related jobs and tourism entrepreneurial migration (Williams and

hall, 2000). In the case of Japan, these kinds of causes might not be strong enough to form a major impact on foreign labour.

# References


[1] Labour Ministry of Health and Welfare. On the registration of information about employing foreign nationals. *http://www.mhlw.go.jp/stf/houdou/0000148933.html*.

[2] Ministry of Internal Affairs Statistic Bureau and Communications. Labor force survey. *http://www.stat.go.jp/data/roudou/sokuhou/nen/ft/index.html*.

[3] OECD UN-DESA. World migration in figures. *A Joint Contribution by UN-DESA and the OECD to the United Nations High-Level Dialogue on Migration and Development*, 2013.

[4] Kazuaki Tezuka. Foreign workers in japan. Technical report, Inter-American Development Bank, 2005.

[5] David Chiavacci. Japan in the global war for talent: Changing concepts of valuable foreign workers and their consequences. *ASIEN*, 124:27–47, 2012.

[6] Gabriele Vogt. Foreign workers in japan. In *The sage handbook of modern Japanese studies*, pages 567–582. Sage London, 2015.

[7] Anna Maria Mayda. International migration: A panel data analysis of the determinants of bilateral flows. *Journal of Population Economics*, 23(4):1249–1274, 2010.

[8] Neelu Seeteram et al. A dynamic panel data analysis of the immigration and tourism nexus. *Monash University Department of Economics, Discussion paper 39*, 10, 2010.

[9] Horst Entorf and Hannes Spengler. Socioeconomic and demographic factors of crime in germany: Evidence from panel data of the german states. *International review of law and economics*, 20(1):75–106, 2000.

[10] Axel Dreher and Panu Poutvaara. Foreign students and migration to the united states. *World Development*, 39(8):1294–1307, 2011.

[11] Cheng Hsiao. *Analysis of panel data*. Number 54. Cambridge university press, 2014.

[12] Luc Anselin. *Spatial econometrics: methods and models*, volume 4. Springer Science & Business Media, 2013.

[13] Jihai Yu, Robert De Jong, and Lung-fei Lee. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics*, 146(1):118–134, 2008.

[14] J Paul Elhorst. Dynamic panels with endogenous interaction effects when t is small. *Regional Science and Urban Economics*, 40(5):272–282, 2010.

[15] Stephen M Renas and Rishi Kumar. The cost of living, labor market opportunities, and the migration decision: A case of misspecification? *The Annals of regional science*, 12(2):95–104, 1978.

[16] Allan M Williams and C Michael Hall. Tourism and migration: new relationships between production and consumption. *Tourism geographies*, 2(1):5–27, 2000.

Table 1: Descriptive Statistics

| Variables (natural logarithm) | Description | Raw data from 2008 to 2013 | |
|---|---|---|---|
| | | Mean | Standard Deviation |
| log(NFW) | Foreign workers (person) | 13423 | 25591.52 |
| log(GPP) | Gross prefecture product (yen) | 10668883 | 14644270 |
| log(CPI) | Consumer Price Index | 99.83 | 1.180484 |
| log(SCU) | Students enrolled in colleges and universities (person) | 54253 | 99764.03 |
| log(RCO) | Recognitions of criminal offenses (person) | 32922 | 42304.42 |
| log(THG) | Total hotel guests (person) | 7238730 | 7045295 |

Table 2: Estimation results

| | Dynamic panel model | Spatial dynamic panel data model | Threshold spatial dynamic panel data model | |
|---|---|---|---|---|
| | | | GPP growth rate $\geq 2.5\%$ | GPP growth rate $< 2.5\%$ |
| $\delta$ (spatial factor) | | 0.2062*** | 0.7403*** | -0.0049 |
| | | (0.0623) | (0.1043) | (0.0809) |
| $\gamma$ (dynamic factor) | 0.4525*** | 0.4268*** | 0.0510 | 0.4787*** |
| | (0.0526) | (0.0458) | (0.1018) | (0.0510) |
| GPP | 1.2758*** | 1.2037*** | 1.1045*** | 1.2470*** |
| | (0.3221) | (0.2389) | (0.4227) | (0.3076) |
| CPI | -1.1068 | -0.1595 | -1.1511 | -0.7657 |
| | (1.2792) | (1.1340) | (1.8947) | (1.3247) |
| SCU | 0.5468 | 0.5712** | 1.7316*** | 0.2649 |
| | (0.3428) | (0.2945) | (0.4574) | (0.3533) |
| RCO | 0.1724** | 0.1906*** | 0.1364 | 0.1382** |
| | (0.0756) | (0.0651) | (0.1853) | (0.0702) |
| THG | 0.1428* | 0.0980 | 0.0840 | 0.1300* |
| | (0.0769) | (0.0677) | (0.1225) | (0.0777) |
| AIC | -594.8720 | -1035.7480 | -1055.274 | |

Note: Numbers in the brackets are standard errors for the corresponding parameter, $*$ shows the result of two-tailed t-test, $*$ indicates the parameter is significant at 10% level, $**$ indicates significance at 5% level, $***$ indicates significance at 1% level.

Table 3: Morans I test for residuals

| year | log(Labor) | residual | | |
|---|---|---|---|---|
| | | Dynamic panel model | Spatial dynamic panel model | Threshold dynamic spatial panel model |
| 2008 | 0.4117*** | | | |
| | (0.0001) | | | |
| 2009 | 0.4194*** | 0.0229 | -0.1275 | -0.2134* |
| | (0.0001) | (0.4104) | (0.3313) | (0.0780) |
| 2010 | 0.4160*** | -0.1316 | -0.2671** | -0.2204* |
| | (0.0001) | (0.3188) | (0.0265) | (0.0706) |
| 2011 | 0.4115*** | 0.2188** | 0.0690 | -0.0270 |
| | (0.0001) | (0.0245) | (0.4010) | (0.9612) |
| 2012 | 0.4026*** | 0.0008 | -0.0887 | -0.0941 |
| | (0.0001) | (0.8383) | (0.5447) | (0.5092) |
| 2013 | 0.3897*** | 0.2215** | 0.0610 | 0.0941 |
| | (0.0002) | (0.0281) | (0.4542) | (0.2929) |

Note: The expected value under the null hypothesis of no spatial autocorrelation is -0.0217 and standard errors for the corresponding Morans I were the numbers in the brackets, $*$ indicates significance at 10% level, $**$ indicates significance at 5% level, $***$ indicates significance at 1% level. log(NFW) is the logarithm of the number of foreign workers, and it is listed for reference.

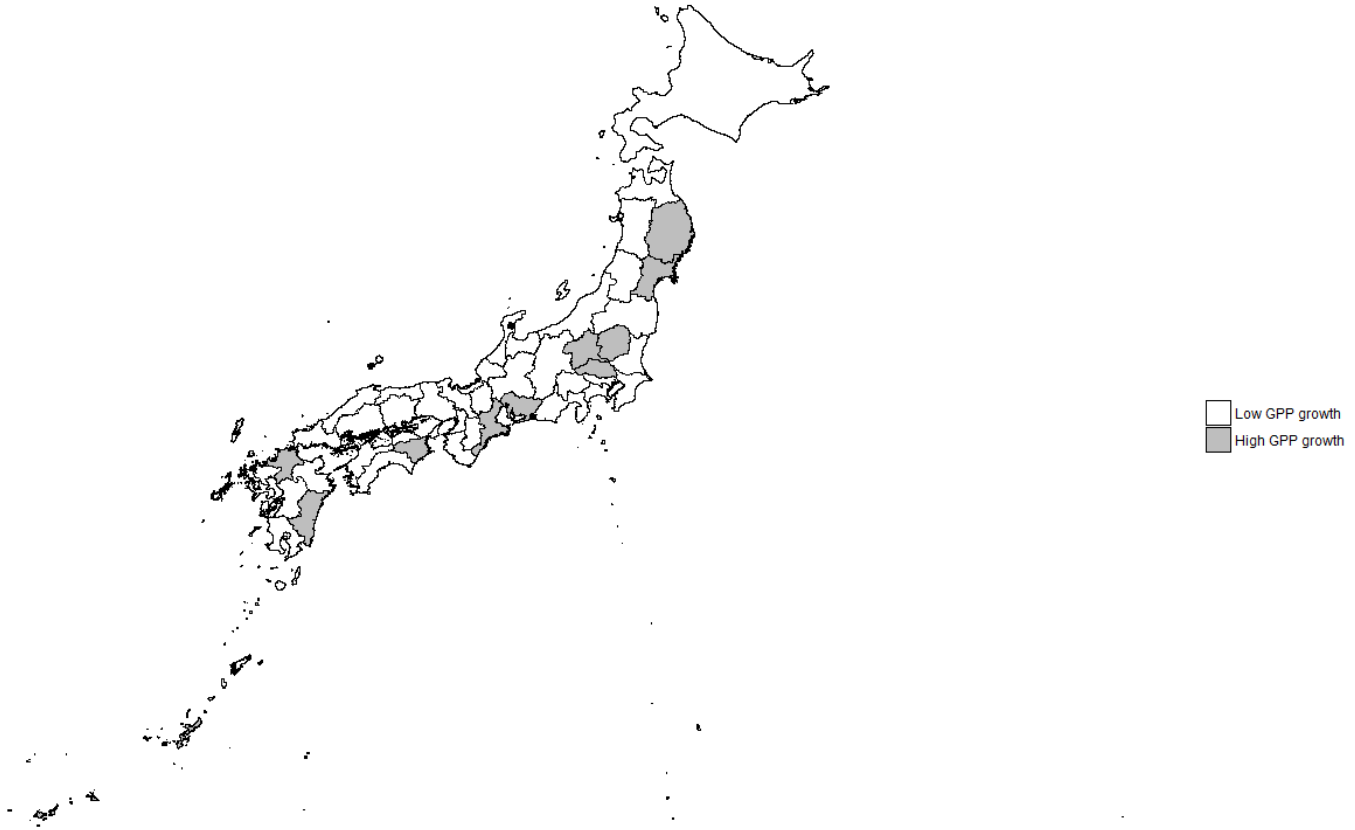



Figure 1: Prefectures divided by growth rate of GPP

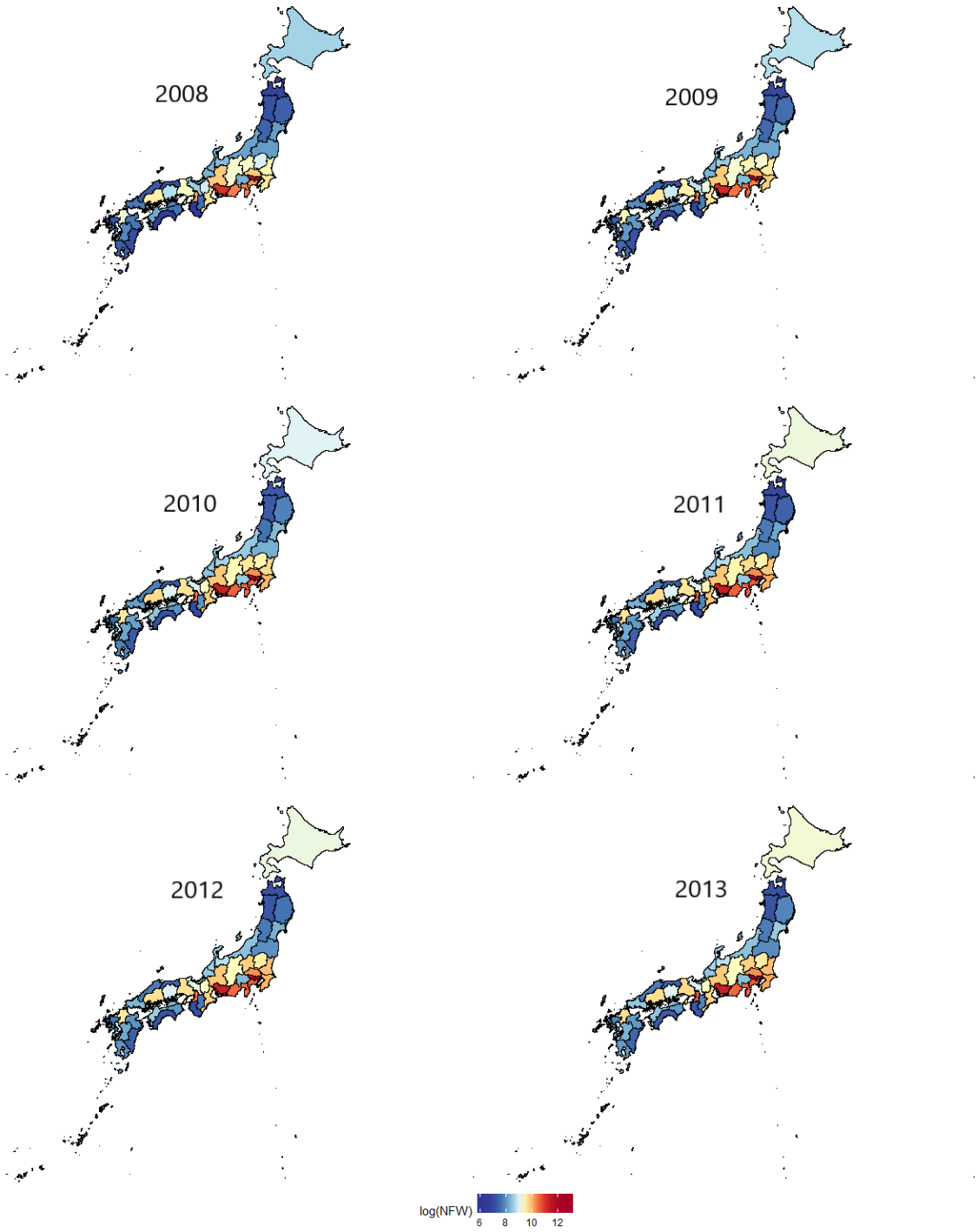



Figure 2: Map chart of foreign workers from 2008 to 2013