Journal of Economics and Development, Vol.17, No.1, April 2015, pp. 5-19 ISSN 1859 0020 | DOI: 10.33301/2015.17.01.01

Using the Spatial Econometric Approach to Analyze Convergence of Labor Productivity at the Provincial Level in Vietnam

Nguyen Khac Minh
Water Resources University, Vietnam
Email: khacminh@gmail.com

Pham Anh Tuan
Vietnam Military Medical University, Vietnam

Nguyen Viet Hung
National Economics University, Vietnam

Received: 13 September 2014 | Revised: 17 January 2015 | Accepted: 08 February 2015

Abstract

This paper employs the spatial econometric approach to undertake a research of labor productivity convergence of the industrial sector among sixty provinces in Vietnam in the period 1998-2011. It is shown that the assumption of the independence among spatial units (provinces in this case) is unrealistic, being in contrast to the evidence of the data reflecting the spatial interaction and the existence of spatial lag and errors. Therefore, neglecting the spatial nature of data can lead to a misspecification of the model. We decompose the sample data into the sub-periods 1998-2002 and 2003-2011 for the analysis. Different tests point out that the spatial lag model is appropriate for the whole period of the sample data (1998-2011) and the sub-period (2003-2011), therefore, we employ the maximum likelihood procedure to estimate the spatial lag model. The estimation results allow us to recognize that the convergence model without a spatial lag variable and using ordinary least square to estimate has the problem of omitting variables, which will have impact on the estimated measure of convergence speed. And this problem dominates the positive effect of factors such as mobilizing factors, trade relation, and knowledge spillover in the regional scope.

Keywords: Spatial econometric; spatial weight matrix; spatial lag model; spatial error model; I-Moran index.

1. Introduction

One hypothesis already proposed by some economic historians, such as Aleksander Gerschenkron (1952) and Moses Abramovitz (1986), is that "following" countries have a tendency to grow more quickly to catch up with the richer ones to narrow the gap between these two groups. This catch up effect is called *convergence.* The question of convergence is central to a lot of empirical research about growth. The neo-classical growth model was built up with the assumption of closed economies. It is derived from the fact that at the beginning, this model is only to explain the progress of growth of one economy. Later, they started using this model to explain the differences in growth rate of per capita income among economies; however, despite these modifications, the original assumption is still kept unchanged, and it is used in empirical analyses about international convergence. William Baumol (1986) is one of the foremost economists providing statistically empirical evidence about the convergence among several countries and the non-existence of convergence among others. Barro and Sala-i-Martin-i-Martin (1991) point out that there is unconditional convergence among states of the US, regions of France, and districts of Japan as we observe in the OECD. The regression method used by Barro has been widely applied in many convergence analyses for different countries such as Koo (1998) considering convergence among regions in Korea, and by Hosono and Toya (2000) considering convergence among provinces in Philippines.

This result is in line with the predictions of the Solow model in the case that provinces within one nation have the same investment rate and population growth rate. However, as we can see, most researches still apply the empirical method for analyzing convergence among countries to the analysis of convergence among provinces within one country. The researchers who mainly pay attention to growth and convergence among regions usually are not aware of the fact that regions and nations are different concepts which cannot be replaced by each other in a simple way.

Although the assumption of a closed economy can be used in an analysis at the international level, it is inappropriate to be applied when analyzing convergence of regions within one country because of much lower restrictions in trade barriers or factor mobilization. Therefore, among many concerns, at least two questions must be emphasized and can suggest a new direction for research: (i) how convergence occurs in the case of an open economy and (ii) how the spatial dependence among regions affects the convergence?

Firstly, if we consider an open economy, we must take the characteristics of factor mobility into account. Factor mobility implies that labor and capital can freely move in response to differences in compensation and interest rates, and they in turn depend on the factor ratios. The capital tends to flow from the regions which have a high capital-labor ratio to the regions which have a lower ratio, and vice versa. In reality, if this adjustment process occurs instantaneously, the speed of convergence approaches infinity.

By bringing the assumption of an imperfect credit market, a finite life-cycle, and the adjustment cost of migration and investment into the model, the speed of convergence to the steady state is finite but larger than the case of a closed economy (Barro and Sala-i-Martin, 1995). Similar results are found when we take trade relations rather than factor mobility into consideration in the neo-classical growth model: the convergence of labor productivity among regions is higher than in the case of a closed economy.

Another possibility for poorer countries to catch up with the richer ones (or having higher labor productivity) is through the spillover effects of technology and knowledge: In the presence of imbalance of technology among regions, the inter-region trade can stimulate a spillover effect of technology when the technological process can be integrated in the tradable commodities (Grossman and Helpman, 1991; Segerstrom, 1991; Barro and Sala-i- Martin, 1997). Another way to explain the spillover effect of technology and knowledge is related to the external effect of knowledge built up by enterprises at a certain location on the production process of other enterprises located in other places. So, the technology spillover effect in the context of productivity convergence implies that the knowledge and technology accumulated, thanks to the spillover effect, can provide opportunities for lagging enterprises (located in low-productivity provinces) to catch up with leading ones (located in high-productivity provinces).

The traditional neoclassical analysis framework can be strengthened by adding the trade relations rather than the flow of factor mobility. Even when there is no factor movement, the balance of prices of tradable goods and the regional specialization based on the relative abundance of factor endowment due to trade can lead to the equalization of factor prices. In addition, when there exists a difference in the level of technology among regions, trade can help enhance the spillover of technology and create opportunities for poorer regions to catch up with richer ones (Nelson and Phelps, 1966; Grossman and Helpman, 1991; Segerstrom, 1991; Barro and Sala-i-Martin, 1995). We can analyze the effect of technology spillover in more detail. Assuming there is no spillover effect of technology, then lagging enterprises cannot catch up with leading ones if they do not invest in R&D or purchase patents to get new technology, however, these are such a huge cost for new entrants into the field as well as for small and medium enterprises. The same argument can be used for differences among regions or provinces. When the spillover effect of technology is not available, the low-productivity provinces cannot catch up with high-productivity ones unless they can invent or buy new technology. However, we should mention that if the spillover effect of technology occurs quickly, one problem can arise. If this effect can occur so easily, then no enterprises have motivation to invest in R&D. In practice, the spillover effect cannot occur immediately but will last for a long period of time. Thereby, the advantage of leading enterprises can be maintained for a certain period of time and helps them to have more incentives to invest into more advanced technology, and convergence only occurs after a while.

In summary, we can expect the speed of convergence to reach the steady state predicted in the version of the neoclassical growth model for an open economy, or in the models with the spillover effect of technology, the speed of con-

vergence would be higher than that in the case of a closed economy.

A direct way to empirically test the prediction of higher speed of convergence for an open economy is to put all variables such as inter-regional movement of capital, labor and technology into the model. However, this direct method has the restriction of the availability of data, especially the data of capital and technology flows as well as technological spillover. A few attempts have been undertaken to test the role of migration flow on convergence. Barro (1991) and Barro and Sala-i-Martin (1995) brought the migration rate as explanatory variables into the regression model for US states, Japanese provinces, and regions of five Asian countries. It is expected that by controlling the migration rate, the estimated speed of convergence would be smaller, and the size of decrease would be a direct measurement of the actual role of migration on speed of convergence. However, in contrast to the authors' expectation, the speed of convergence was almost always not affected by putting this variable into the model, even when we use the instrumental variable to take the possibly endogenous effect on migration rate into account. These results, together with the fact that the net migration rate tends to positively respond to the initial level of per capita income, advocate for the view that migration has little effect on speed of convergence, whereas most of the effect on this process comes from the change in capital-labor ratio, which is determined by saving rate.

In summary, the neoclassical model describes a tendency of the whole economy system. It approaches not only to the equilibrium of the market in markets of each region but also the general equilibrium in the inter-connection between each region and the rest of the whole system. These regions build up a system, as described by the authors, including residents sharing a similar technology system. This implies that these regions would have the same steady state. Therefore, in such a framework, differences in economic growth of regions are mainly due to two causes: (i) growth of capital stock per capita is financed by internal resource, and (ii) a quick decrease in the initial misallocation of resources among regions thanks to the openness of the region. Combining these two factors, the speed of convergence to the steady state would occur more quickly than in the case of a closed economy. After understanding the important role of the mobility among regions due to their openness in explaining the regional convergence, now we can continue to study the spatial interaction effect on the convergence analysis from the econometric perspective.

In general, two main causes of misspecification which have been pointed out in research on spatial econometric are: (i) spatial dependence and (ii) spatial heterogeneity (Anselin, 1988). Spatial dependence (or spatial autocorrelation) originates from the dependence of observations ranked by the order of space (Cliff and Ord, 1973). Specifically, Anselin and Rey (1991) distinguish between strong and disturbance spatial dependence. Strong spatial dependence reflects the existence of the spatial interaction effect, for instance, the spillover effect of technology or the mobility of factors, and these are the crucial components determining the level of income inequality across regions. Disturbance spatial dependence can originate from troubles in measurement such as the incompatibility between spatial features in our research and the spatial boundary of observation units. The second cause of misspecification, i.e. spatial heterogeneity, reflects the uncertainty of the behavioral aspects among observation units.

As Rey and Montuori (1999) emphasized, researches of spatial econometrics have provided a series of procedures to test the existence of the spatial effect (Anselin, 1988; Anselin, 1995; Anselin and Berra, 1998; Anselin and Florax, 1995; Getis and Ord, 1992). Additionally, in the cross-section approach, there are some forms of estimation parameters for models explicitly considering spatial effects. The version of strong dependence to study spatial dependence is called as spatial autocorrelation model (Anselin and Bera, 1998; Arbia, 2005), or spatial lag model. Some empirical researches have used the econometric background to test the regional convergence. The most complete researches which can be mentioned include Rey and Montouri (1999), Niebuhr (2001), and Le Gallo et al. (2003) and Abria and Basile (2005).

This research includes four sections. The next section presents the background of methodology including this content: how to construct a weight matrix, spatial lag models, a spatial error model, and some important tests. The third section briefly describes the data and estimation results. Finally, the conclusion is given in the fourth section.

2. Theoretical framework

2.1. Method to identify the weight matrix

To study spatial convergence, we have to construct the model and test the existence of spatial dependence. To develop the model, we need to construct the weight matrix and do some necessary tests. Hence, in this section, we present how to identify a weight matrix w.

The spatial econometric model which we will build up will use provinces as the spatial units. Normally, in empirical analyses, administrative units are most popularly used. In the context of Vietnam, taking provinces as the spatial units is the most appropriate because the data at the provincial level are available. The method to identify a weight matrix is as follows: For each province, we identify a central point (the city or the town). We can identify the latitude and longitude of this central point by using a geographical map. Using the Euclidian distance in the two-dimension space, we have:

$$d_{ij} = d\left(s_i, s_j\right) = \sqrt{\left(s_i - s_j\right)^T \left(s_i - s_j\right)} \quad (1)$$

In which d_{ij} is the distance between two points s_i and s_j. Two provinces would be called neighbors if $0 \le d_{ij} < d^*$, d_{ij} is the distance which is computed by using the formula (1), d^* is called the critical cutting point. We also define two provinces i and j to be called as t neighbors if $d_{ij} = \min\left(d_{ik}\right), \forall i, k$. Denote N(i) as the collection of all neighbors of province i. Then, the binomial weight matrix is the matrix with elements identified as follows:

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

Denote $\eta_j = \sum_i w_{ij}$, and $w_{ij}^* = w_{ij}/n_j$, then $W^* = \left[w_{ij}^*\right]_{n \times n}$ is called a row-standardized binary version of a spatial weight matrix. Using this methodology, we can construct the weight matrix for the productivity convergence model of sixty provinces (sixty spatial units in the empirical research).

2.2. β- convergence

So far, the β-convergence approach is still considered as the most persuasive theoretical approach from the economic theory perspective. At the aspect of policy making, this is also a highly persuasive approach because it can identify an important concept relating to speed of convergence. It can go beyond the neoclassical growth model of Solow-Swan, in which it is assumed that the economy is closed, the saving rate is endogenous, and the production function has the features of decreasing returns with respect to capital and a constant return to scale. This model predicts that the growth rate of a region is positively correlated to the distance from the current position of the economy to its steady state. Some authors such as Mankiw et al. (1992) and Barro and Sala-i-Martin (1992) suggested a statistical model using cross-section units in the form of a matrix as follows:

$$\frac{1}{T}\ln\left[\frac{y_T}{y_0}\right] = \mu_{0,T} + \varepsilon \quad (2)$$

$$\varepsilon \sim \aleph\left(0, \sigma_\varepsilon^2 I\right)$$

In which y_T is the value of labor productivity on average at the end point of the period under consideration, y_0 is the value of the first period and ε is the identically and independently distributed error component (i.i.d) and it is the unsymmetrical component of the model. $\mu_{0,T}$ is the symmetrical component of the model and is identified as follows:

$$\mu_{0,T} = \alpha' - \frac{\left(1-e^{-\lambda T}\right)}{T}\ln y_0 \quad (3)$$

In which, λ is the speed of convergence, measuring the speed at which the economy will converge to its steady state. From the model (2) and (3), we can get this model:

$$\frac{1}{T}\ln\left[\frac{y_T}{y_0}\right] = \alpha' - \frac{\left(1-e^{-\lambda T}\right)}{T}\ln y_0 + \varepsilon \quad (4)$$

The model (4) is normally directly estimated by using Non-Linear Least Square (Barro and Sala-i-Martin, 1995), or statistical model - parameterized by letting $\beta = -(1 - e^{-\lambda T})$, $\alpha = T\alpha'$, $\lambda = -\ln(1+\beta)/T$, the model (4) becomes:

$$\ln\left[\frac{y_T}{y_0}\right] = \alpha + \beta\ln y_0 + \varepsilon \quad (5)$$

Then, β can be estimated by using the ordinary least square method. The absolute convergence exists when the estimation of β takes the negative value and is statistically significant. If the null hypothesis $(\beta=0)$ is rejected, then we can conclude that not only the regions which have lower productivity will grow more quickly, but all of them will converge to the same level of labor productivity.

The constant component, α depends on y^*, in which y^* is labor productivity at the steady state. In these settings, all provinces are assumed to be homogeneous in terms of structure and can have access to the same type of technology, so they can be characterized by the same steady state, and the only difference among these economies are the initial conditions.

In the scope of this paper, the concept of β conditional convergence will be employed when the assumption of the same steady state is relaxed.

2.3. Moran index

The σ-convergence approach is to compute the standard error of per capita income of regions and to analyze the long-term tendency of this value. If this value tends to decrease, regions will converge to the same level of income. In this approach, a problem arising is that the standard deviation is very difficult to be recognized for spatial units, and it does not allow to distinguish between very different geographical conditions (Arbia, 2005). Moreover, according to Rey and Montouri (1999), the σ-convergence analysis can "*veil the unusual geographical forms which can vary overtime*". Therefore, it is useful to analyze geographically spatial dimensions of income distribution together with dynamic behavior of income variations. This is quite possible by using I-Moran statistics to examine different forms of spatial autocorrelation (Cliff and Ord, 1973). The I-Moran test statistic can be identified as follows:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \hat{e}_i \hat{e}_j}{\sum_{i=1}^{n} \hat{e}_i^2} \quad (6)$$

In which $\hat{e}_i = y_i - \hat{\beta}^T x_i$ is the residuals of OLS estimation, $w_{ij} \in W$, W is the binominal spatial weight matrix. Written in the form of a matrix, the formula (6) then becomes:

$$I = h\left(\hat{e}^T \hat{e}\right)^{-1} \left(\hat{e}^T W \hat{e}\right) \quad (7)$$

In which $\hat{e} = y - \hat{\beta}^T X$ and X are data matrix. If we employ the row-standardized binomial weight matrix, then

$$I = \left(\hat{e}^T \hat{e}\right)^{-1} \hat{e}^T W^* \hat{e} \quad (8)$$

Because the residuals follow the normal distribution, then the I-statistic approaches the normal distribution, in which the expectation value is

$$E(I) = \frac{k-1}{n-k-1} tr\left(MW^*\right)$$

and variance is

$$V(I) = \frac{tr\left(MW^* MW^{*T}\right) + tr\left(MW^*\right)^2 + \left[tr\left(MW^*\right)\right]^2}{(n-k-1)(n-k+1)} - E(I)^2$$

In which, $M = I - X(X^T X)^{-1} X^T$. The positive and significant value of I-Moran implies spatial convergence while the negative value implies spatial divergence.

2.4. Spatial dependence in the cross-section growth equation

The neoclassical growth model mentioned above has been developed on the basis of a closed economy. However, this assumption is so strong for the analysis of regions within one country, in which there exists negligible trade and factor mobility barriers (Magrini, 2003). To understand implications of bringing the assumption of an open economy into the model with respect to convergence, we must consider the role of factor mobility, trade relations and the spillover effect of technology or knowledge.

After clarifying the important role of mobility flows across regions due to their openness on regional convergence, now we can turn to the second question that we have mentioned above, and we examine the effects of spatial interaction on convergence analysis from the econometric perspective.

In general, two main causes of misspecification which have been pointed out in researches of spatial econometric are (i) spatial dependence and (ii) spatial heterogeneity (Anselin, 1988). Spatial dependence (or spatial autocorrelation) originates from the dependence of observations ranked by the order of space (Cliff and Ord, 1973). Specifically, Anselin and Rey (1991) distinguish between strong and disturbance spatial dependence. The strong spatial dependence reflects the existence of a spatial interaction effect, for instance, the spillover effect of technology or the mobility of factors, and they are the crucial components determining the level of income inequality across regions. Disturbance spatial dependence can originate from troubles in measurement, such as the incompatibility between spatial features in our research and the spatial boundary of observation units. The second cause of misspecification, i.e. spatial heterogeneity, reflects the uncertainty of the behavioral aspects among observation units.

The first strong dependence form can be integrated into the traditional cross-section specification by the spatial lag of the dependent variable, or spatial lag model. If W is the row-standardized spatial weight matrix which describes the structure and intensity of the spatial effect, then the spatial lag model has the following form:

$$g_{\ddot{u},} = \ln\left[\frac{y_{T,i}}{y_{0,i}}\right] = \alpha + \beta\ddot{u}\;_{0i} + \rho\sum_{j=1}^{n} \mathrm{w_{i,j}} \ln\left[\frac{y_{T,i}}{y_{0,i}}\right] + \varepsilon \quad (9)$$

In which ρ is the parameter of the spatial lag dependent variable, $\sum_{j=1}^{n} \mathrm{w_{i,j}} \ln\left[\frac{y_{T,i}}{y_{0,i}}\right]$ captures the interaction impact, showing how the growth rate of GDP per capita in one region is determined by the growth rate in neighboring regions. The error component is assumed to be identically, independently and normally distributed (i.i.d) and it is assumed that all spatial dependence effects are consisted in the lag component.

The specification (4) can be written in the vector version as follows:

$$g_T = \ln\left[\frac{y_T}{y_0}\right] = \alpha + \beta lny_0 + \rho \mathrm{w} \ln\left[\frac{y_T}{y_0}\right] + \varepsilon \quad (10)$$

Putting the term ρwln(y_T/y_0) to the left-side, we have

$$(1-\rho \mathrm{w})\ln\left[\frac{y_T}{y_0}\right] = \alpha + \beta lny_0 + \varepsilon \quad (11)$$

The model (11) can be interpreted in different ways but the most important is the nature of convergence after controlling the effect of spatial lag.

The parameters in model (10) can be estimated by the maximum likelihood method (ML), instrumental variables, or procedures of general moment method.

Now, we can specify the spatial lag model.

We can integrate the spatial effects through the spatial error model which has been proposed by Anselin and Bera (1998), Arbia (2005). Using vector denotation, the errors can be identified as follows:

$\varepsilon_t = \psi \mathrm{W} \varepsilon_t + u_t$,

Moving the first term of the right-side to the left-side of the equation, we have:

$\varepsilon_t = (I - \psi \mathrm{W})^{-1} + u_t$

In which ψ is the coefficient of spatial error and $u \sim N(0,\sigma^2 I)$. In this case, the original error has the covariance matrix in the form of a non-spherical form:

$E[\varepsilon\varepsilon'] = (I - \psi W)^{-1}\sigma^2 I(I - \psi W)^{-1}$

So, using the ordinary least square method (OLS) in the presence of non-sphere error would make the estimation of convergence parameter bias. As a consequence, the OLS applied for the spatial lag model would provide inconsistent estimations, and we should employ estimations based on the maximum likelihood and instrumental variable method (Anselin, 1988). From the spatial analysis perspective, an interesting feature of the disturbance dependence model has been clarified in Rey and Montuori (1999). In this case, a random shock which has effect on a certain region will have effect on the growth rate of other regions through the spatial variation component. In other words, any movements that diverge from the growth pattern of the steady state may not only depend on the shock characterized by regions, but also depend on the spillover effect of shocks from other regions.

2.5. A test of spatial dependence

As Rey and Montuori (1999) emphasize, researches of spatial econometrics have provided a series of procedures to test the existence of the spatial effect (Anselin, 1988; Anselin, 1995; Anselin and Berra, 1998; Anselin and Florax, 1995; Getis and Ord, 1992). The tests, based on two types of econometric model, namely the spatial lag model and the spatial error model, can be in the form of the Lagrange multiplier test (LM), and the test suggested by Anselin et al. (1996) which uses the Monte Carlo method to examine a finite sample and a trend test to provide the correction method for the LM test to test the spatial dependence characteristic. They found that the corrected LM method for a finite sample has many attributes. This paper employs the LM test method suggested by Anselin (1995) to select the more appropriate model.

A test of the existence of spatial autocorrelation errors

H_0: non-existence of spatial dependence (spatial autocorrelation) (H_0: σ=0)

The test statistic:

$$LM_{error} = \left(\frac{e'\mathrm{w}e}{e'e} \right) / \, tr\left(\mathrm{w}'\mathrm{w} + \mathrm{w}^2 \right)$$

In which tr is the matrix trace; e is the vector of OLS residuals; W is the row-standardized spatial weight matrix.

The LM statistic follows the $\chi^2(1)$ distribution.

A test of the existence of spatial lag

H_0: non-existence of spatial lag dependence (H_0: ρ=0)

The test statistic:

$$LM_{Lag} = \left(\frac{e'\mathrm{w}_g}{e'e} \right) \left(\mathrm{w} \ln y_0 b \right) / \, e'e + tr\left(\mathrm{w}'\mathrm{w} + \mathrm{w}^2 \right)$$

In which w_g is the spatial lag of the dependence variable; b is the least square estimation of the parameter β. The LM statistic follows the $\chi^2(1)$ distribution.

3. Empirical results

3.1. Data

The objective of this paper is to analyze the convergence of the labor productivity of the whole economy and three economic sectors including agriculture, industry, and services at the provincial level. The data, including output, capital, and labor compensation in the period 1998-2011 are collected from the General Statistical Office, Ministry of Labor, Invalids and Social Affairs. This data set consists of the output computed at constant prices, the net value of capital at a constant price, and the labor of the whole economy and of three sectors.

However, there exists one problem with this data set. Firstly, due to the merging and splitting of provinces, some provinces are available only in some years in this period. To guarantee the pureness of research units, we decide to aggregate the data of some provinces as follows: combining the data of Hanoi and Ha Tay, Dak Lak and Dak Nong, Dien Bien and Lai Chau, Can Tho and Hau Giang.

In an analysis of convergence, the central issue is the relative value of labor productivity because we want to see if the provinces with low-productivity can grow more quickly than the ones with high-productivity. This data set is not biased due to sample selection (because all provinces are brought into the analysis), and we can expect that the relative growth of provinces are compatible.

At first, we employ cross-section regression to estimate the convergence of labor productivity for the whole economy, and estimate labor productivity convergence at the provincial level of three sectors, namely agriculture, industry and service. It is shown that the estimation results do not support for the hypothesis of convergence of labor productivity in the case of the agriculture sector and the whole economy.

We employ the spatial econometric techniques to estimate labor productivity convergence in sixty provinces for two sectors: industry and service. We find out that the econometric model used for the service sector does not satisfy some tests, therefore, in the following section, only the estimation results of the labor productivity convergence for the industry sector would be provided.

3.2. Empirical results

Table 1 gives the estimation results using the ordinary least square method for the case of unconditional convergence of labor productivity in the industry sector in sixty Vietnamese provinces in the whole period 1998-2011 and two sub-periods (1998-2002 and 2003-2011).

In this model, the dependent variable expresses the growth rate of labor productivity on average in the whole period and two sub-periods. The OLS estimation coefficient of the initial labor productivity for the whole period is highly statistically significant and takes a negative value. This confirms the existence of the absolute convergence of labor productivity in the industry sector in the period 1998-2011. When we decompose the whole period into two sub-periods, 1998-2002 and 2003-2011, the estimation results give us interesting insights. There is evidence about the different patterns in the growth of labor productivity in the provinces. The coefficients of the initial labor productivity for the two sub-periods are respectively -0.2623 and -0.3969, and both of them are statistically significant.

Table 1 also provides the results of different model specification tests based on the cross-section data and the residuals from the OLS estimation. The value of the Jarque-Bera test is not significant, implying that the null hypothesis, errors following the standard distribution, is not rejected. So, we can explain that the results of the misspecification test (the heterogeneity of variance test, spatial dependence test) are meaningful. The value of the Breusch-Pagan test statistic shows that there is no variance heterogeneity, except the model in the period 1998-2002. The result of this test is once again affirmed by the White test. Table 1 also gives the result of the maximum likelihood function and value Schwartz and AIC criterion. These criteria imply that the convergence model estimated by the OLS technique for the whole period and the second sub-period are

Table 1: The estimation results of unconditional convergence of labor productivity in the industry sector using OLS

	1998-2011	1998-2002	2003-2011
α	2.418844	.8189173	1.890344
	(0.000)	(0.000)	(0.000)
β	-.5596322	-.2623358	-.3968612
	(0.000)	(0.000)	(0.000)
Goodness of fit			
Adjusted R^2	0.3935	0.2208	0.2210
Log likelihood	-41.87277166	-20.29362622	-36.34453791
AIC	1.462426	.7431209	1.278151
Schwartz Criterion	-223.2864	-230.5623	-225.6737
Regression Diagnostic			
Jarque-Bera	.0914	2.342	.5446
	(.9553)	(.3101)	(.7616)
Breusch-Pagan	2.007446	9.721399	.8667575
	(.5709)	(.0211)	(.8334)
White	.0406363	12.47832	0.9555399
	(0.8402)	(0.0004)	(0.3283)
Moran's I	1.866	-0.950	2.300
	(0.062)	(1.658)	(0.021)
LMe	1.278	1.432	2.218
	(0.258)	(0.231)	(0.136)
Robust LMe	1.143	0.014	1.847
	(0.285)	(0.904)	(0.174)
LM Lag	3.761	1.779	4.494
	(0.052)	(0.182)	(0.034)
Robust LM Lag	3.626	0.361	4.123
	(0.057)	(0.548)	(0.042)

Source: The author's estimation using the data set of General Statistics Office of Vietnam (GSO) and Ministry of Labour - Invalids and Social Affairs (MOLISA).

Note: The number in parentheses is the probability.

approximate to each other (AIC in the whole period model is 1,4624 whereas its value in the second sub-period model is 1,2781).

There are three different tests for the existence of spatial dependence. They are Moran I, and two Lagrange multiplier tests. The first test shows that the null hypothesis is rejected at the 10% significance level for the whole period and at the 5% significance level for the second sub-period. This is a powerful test, however it does not allow us to identify the cause of misspecification as a consequence of spatial lag or spatial errors (Anselin and Rey, 1991). Table 1 also provides the results of the two Lagrange multiplier tests (LM), in which the test of spatial error is not significant in any period under consideration while the Lagrange multiplier test of spatial lag is significant at the 10% level for the whole period and 5% for the sub-period 2003-2011.

Table 2: Estimation results of labor productivity in Vietnam using the spatial lag model and maximum likelihood method

	1998-2011	1998-2002	2003-2011
α	2.050751	.932069	1.509757
	(0.000)	(0.000)	(0.0)
β	-.5419362	-.2591993	-.3718775
	(0.000)	(0.000)	(0.0000)
ρ	.3143677	-.7400394	.378562
	(0.152)	(0.109)	(0.104)
Adjusted R^2	0.8374	0.2904	0.7810
Log-Likelihood	-40.923719	-18.96498	-35.170278
AIC	0.2540	0.1357	0.2193
Schwartz Criterion	0.2724	0.1455	0.2352
Spatial Breusch-Pagan heteroschedasticity test	0.0001	0.2107	0.4891
	(0.9912)	(0.6462)	(0.4843)
LR test spatial autocorrelation	1.898	2.657	2.349
	(0.168)	(0.103)	(0.125)
LM test(error)	3.761	1.779	4.494
	(0.052)	(0.182)	(0.034)

Source: the author's estimation using the data set of GSO and MOLISA.

In summary, the least square estimation of the convergence model is misspecified due to the effect of spatial lag, i.e. the labor productivity of each province is not independent of the other provinces' labor productivity.

According to the above tests, the spatial lag model is suitable for the whole period (1998-2011) the second sub-period (2003-2011). Therefore, we would use the maximum likelihood procedure to estimate the spatial lag model. The results are given in Table 2.

Table 2 gives the results of the spatial lag model estimated by the maximum likelihood method (ML). The estimate parameters are highly statistically significant. We can compare the coefficient of logarithm of the labor productivity estimated by OLS and the one estimated in the spatial lag model by the maximum likelihood method in the whole period and the second sub-period 2003-2011. The coefficient estimated by OLS, and not taking the effect of spatial lag into consideration in the whole period and the sub-period 2003-2011, are respectively -0,5596 and -0,3968. Meanwhile, the coefficient in the spatial lag model estimated by the maximum likelihood method in these two periods are respectively -0,5419 and -0,3719. Comparing these results shows that the coefficients of logarithm of the labor productivity in the spatial lag model are smaller in absolute value in both periods. The decrease in the value of these coefficients is due to the presence of the spatial lag effect in the model. The economic reasons for this characteristic can be explained as follows. Firstly, it is the effect of omitting a variable, i.e. putting the spatial lag variable into the model can help correct the model in terms of spatial dependence. The representative variable for the spatial dependence

Table 3: Comparing the speed of convergence and half-life time in the two periods

	1998-2011	2003-2011
Spatial Lag Convergence Rate Estimated	0.060057	0.058128
Spatial Lag Half-Life	11.54	11.92
OLS Convergence Rate Estimated	0.063088	0.063201
OLS Half-Life	10.98	10.96

Source: the author's estimation using the data set of GSO and MOLISA.

can capture the effects of variable omission (the difference comes from migration, trade, and spillover effect. The variable omission can have a negative effect on the growth of productivity. Secondly, it is the positive effect of factor mobility (labor mobility across provinces for instance), trade relations, and knowledge spillover at the regional level. The technology and knowledge spillovers have an important role. The technology spillover behind the productivity convergence can bring about opportunities for enterprises in lagging provinces to catch up with leading enterprises. Assume that there is no technology spillover. Then, lagging enterprises cannot catch up with leading ones if they do not invest in R&D or purchase patents to get new technology, however, these present such a huge cost for new entrants into the field as well as for small and medium enterprises. The same argument can be used for differences among regions or provinces. When the spillover effects of technology are not available, the low-productivity provinces cannot catch up with the high-productivity ones unless they can invent or buy new technology. However, we should mention that if the spillover effect of technology occurs quickly, one problem can arise. If this effect can occur so easily, then no enterprises have motivation to invest in R&D. In practice, the spillover effect cannot occur immediately but it lasts for a long period of time. Thereby, the advantage of leading enterprises can be maintained for a certain period of time and helps them to have more incentives to invest into more advanced technology, and convergence only occurs after a while. However, the sum of these two effects can be negative or positive, depending on which effect dominates. Table 3 compares the speed of convergence and half-life time estimated in the spatial lag model for two periods.

The estimation and test results in Table 1 and 2 show that there exists a spatial lag effect, i.e. if there is no other effect, the positive effect of spatial lag effect would make the speed of convergence increase as in the theoretical explanation above. However, looking at the results in Table 3 shows that the speeds of convergence in the spatial lag model are 6% for the whole period and 5,8% for the sub-period 2003-2011, while they are 6,3% and 6,32% in the model without spatial lag effect. These results are opposite to what the theory explains. However, this estimation result helps us find out that the versions of convergence model suggested by Mankiw et al (1992) and Barro and Sala-i-Martin (1992) have the problem of variable omission. The omission of the variable has a negative impact on the speed of convergence. This omission effect dominates the positive effect of factors such as factor mobility, trade relations, and knowledge spillover at the regional level. And that explains why the speed of convergence in the spatial lag model is less than that in the traditional model.

4. Conclusion

We have studied the convergence of labor productivity in the industry sector in sixty provinces in Vietnam in the period 1998-2011 by employing the spatial econometric approach. Two issues are discussed in this paper: how does the spatial dependence among regions affect the convergence. In general, two causes of misspecification have been pointed out in researches of spatial econometric: spatial dependence and spatial heterogeneity. We employ the spatial econometric approach to estimate the model. We point out that the least square estimation of the convergence model causes misspecification due to the existence of spatial lag in the model, i.e. the labor productivity in each province is not independent of the others. The estimation results show that there exists a spatial lag effect, however the impact of variable omission dominates the positive effect of factor mobility, trade relations, and knowledge spillover at the regional level.

Acknowledgements

Funding from Vietnam National Foundation for Science and Technology Department (NAFOSTED), under grant number II 2.2-2012.18 is gratefully acknowledged.

We would like to send special thank to the editor and two anonymous referees.

References

Abramovitz M. (1986), 'Catching up Forging Ahead and Falling Behind', *Journal of Economic History*, Vol. 46, pp. 385-406.

Anselin L. and Bera A.K. (1998), 'Spatial dependence in linear regression models with an introduction to spatial econometrics', in Hullah A. and D.E.A. Gelis (eds.) *Handbook of Applied Economic Statistics*, Marcel Deker, New York, pp. 237-290.

Anselin L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer Academic Publishers, Dordrecht.

Anselin L. and Rey S.J. (1991), 'Properties of test for spatial dependence', *Geographical Analysis*, Vol. 23, pp. 112-131.

Anselin L., Bera A., Florax R., and Yoon M. J. (1996), *Simple diagnostic tests for spatial dependence*, IDEAS.

Anselin, L. (1995), 'Local Indicators of Spatial Association – LISA', *Geographical Analysis*, Vol. 27, Issue 2, pp. 93-115.

Anselin, L. and Florax, R. (Eds.) (1995), *New Directions in Spatial Econometrics*, Berlin: Springer.

Arbia G. and Basile R. (2005), *Spatial Dependence and Non-linearities in Regional Growth Behaviour in Italy*, Statistica.

Barro R.J. and Sala-i-Martin X. (1992), 'Convergence', *Journal of Political Economy*, Vol. 100, pp. 223-251.

Barro R.J. and Sala-i-Martin X. (1995), *Economic Growth*, McGraw Hill, New York.

Barro R.J. and Sala-i-Martin X. (1997), 'Technological diffusion, convergence and growth', *Journal of Economic Growth*, Vol. 2, pp. 1-16.

Barro, Robert J. and Xavier Sala-i-Martin (1991), 'Convergence Across States and Regions', *Brookings Papers on Economic Activity*, pp. 107-182.

Baumol, William J. (1986), 'Productivity Growth, Convergence, and Welfare: What the LongRun Data Show', *American Economic Review*, Vol. 76, pp.1072-1085.

Cliff, A. and Ord, J. (1973), *Spatial Autocorrelation*, London: Pion.

Gerschenkron, A. (1952), 'Economic Backwardness in Historical Perspective', In The *Progress of Underdeveloped Areas*, ed. Bert F.Hoselitz, Chicago: University of Chicago Press.

Getis, A., and J. K. Ord (1992), 'The Analysis of Spatial Association by Use of Distance Statistics', *Geogmphicd Analysis*, Vol. 24 (July), pp. 189-206.

Grossman, G. M. and Helpman, E. (1991), *Innovation and Growth in the Global Economy*, Cambridge MA: MIT Press.

Hosono, Kaoru. and Hideki Toya (2000), 'Regional Income Convergence in the Philippines', Discussion Paper No. D99-22, Institute of Economic Research, Hitotsubashi University, February 2000.

Koo, Bon Cheon. (1998), *The Economic Analysis of Corporate Exit and the Reform Proposals (in Korean)*, Seoul: Korea Development Institute.

Le Gallo J., Ertur C., and Baoumont C. (2003), 'A spatial Econometric Analysis of Convergence Across European Regions, 1980-1995', in Fingleton, B (ed.), *European Regional Growth*, Springer-Verlag (Advances in Spatial Sciences), Berlin.

Magrini S. (2003), 'Regional (Di)Convergence', in V. Henderson and J.F., Thisse (eds.), *Handbook of Regional and Urban Economics*, Volume 4.

Mankiw N.G., Romer D. and Weil D.N. (1992), 'A contribution to the empirics of economic growth', *Quarterly Journal of economic*, Vol. 107, pp. 407-437.

Nelson, R. R. and Phelps, E. S. (1966), 'Investment in Humans, Technological Diffusion, and Economic Growth', *American Economic Review*, Vol. 56, No.2, pp. 69-75.

Niebuhr, A. (2001), 'Convergence and the Effects of Spatial Interaction', *Jahrbuch für Regionalwissenschaft*, Vol. 21, No. 2, pp. 113-133.

Rey S. J. and Montuori, B. D. (1999), 'US Regional Income Convergence: A Spatial Econometric Perspective', *Regional Studies*, Vol. 33, No.2, pp. 143-156.

Segerstrom, P. S. (1991), 'Innovation, Imitation, and Economic Growth', *Journal of Political Economy*, Vol. 99, No.4, pp. 807-827.