Spatial Econometric Research on the Relationship between Highway Construction and Regional Economic Growth in China: Evidence from the Nationwide Panel Data

N J Ye, W J Li, Y Li and Y F Bai

1 School of Economics and Management, Chang'an University, Xi'an 710064, China
2 School of Mathematics and Computational Science, Xiangtan University, Xiangtan 411100, China
Email: serinalee150@gmail.com

Abstract. Based on spatial panel data from 2010 to 2016 in China, this paper makes an empirical analysis on the relationship between highway construction and regional economic growth by means of spatial econometric model. The results show that there is positive spatial correlation on regional economic growth in China, and strong spatial dependences between some provinces and cities appear, specifically, Hebei, Beijing, Tianjin, Shanghai, Zhejiang and other eastern coastal areas show high-high agglomeration trend; the Pearl River Delta region presents high-low agglomeration trend; In terms of nationwide provinces and municipalities, a province's highway construction investment for their own province and the neighboring provinces has pulling effect on economic growth to a certain extent, and the direct effect is more obvious.

1. Introduction

As a transportation infrastructure, highway construction has always been considered to be one of the important factors to promote regional economic growth in China. The study on the relationship between highway construction and regional economic growth has been a hot topic for domestic and foreign scholars. Most scholars believe that highway construction has influence on regional economic growth in a way (Ye C Y et al., 2013; Zheng G J, 2014) [1][2]. About research methods, it’s mainly divided into qualitative analysis and traditional quantitative analysis, such as inductive analysis, deductive analysis, input-output model, time series analysis, regression analysis and so on. In the past, the relationship between highway construction and regional economic growth has been deeply explored, but there are still some aspects for improvement on the research methods. First, the qualitative analysis is susceptible to the influence of various subjective factors, such as knowledge level and practice degree, and the phenomenon of multiple collinearity exists in the analysis on input and output model. Second, based on the analysis results of traditional cross-sectional data and pure panel data, there is the possibility of exaggerating the effect on highway construction in promoting regional economic growth. Simultaneously, many scholars pointed out that, for China, spatial dependencies and heterogeneities on economic development between different provinces and cities were obvious. Zhang X L conducted a spatial empirical analysis on the relationship between transport infrastructure and regional economic growth in 2012, it was found that the former had an effect of strong spatial spillover on the latter, and the factor of new economic geography had played a better role in regional economic growth[3]. The scholar, Yu J X, analyzed the spatial dependence and
heterogeneity between highway construction and nationwide, four economic regions in 2015, the result showed the spatial differences on highway construction in China were significant, the change trend of highway construction and regional economic were proved with same direction. In addition, the author, Chu C L, confirmed the spatial effect of highway construction on regional economic growth was a dynamic development process with forward-negative-equilibrium in 2015. In 2016, a spatial correlation study on the relationship between them was conducted by Li Q Z, being proved highway construction had positive effect on regional economic growth, and the role of expressway for promotion was more obvious.

Spatial econometric model are welcomed by many scholars with rapid development of regional economic. Based on the analysis of spatial panel data, it can effectively solve the problem of multiple collinearity of the model, and the characteristics of spatial differences and dependency in different provinces and cities are considered comprehensively. Therefore, on the basis of the spatial panel data of 31 provinces and cities in China from 2010 and 2016, this paper makes a deep research on the relationship between highway construction and regional economic growth by using spatial econometric model for improving research framework existing and promoting the study on econometric method further.

2. Theoretical model

2.1 Spatial statistical analysis

2.1.1 Global spatial autocorrelation test
Global spatial autocorrelation is used for describing the situation of different predictive values in different spaces and reflecting the state of spatial aggregation and discretization of variables. Moran's I index and Geary C index are the main indicators for the global spatial autocorrelation test, in which Moran's I index is widely used, as shown in equation (1):

\[
Moran's I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \sum_{j=1}^{n} (X_j - \bar{X})^2}
\]  

In the formula (1), “n” means the total number of the study areas, “W” means spatial weight matrix, “X” represents the observed value of the study area of “I”. Moran's I value is in the range of [-1,1], and when the value belongs to [-1,0), it indicates that the global space is negatively correlated, and dissimilar observed values have agglomeration effect, being more significant when values being closer to -1. When the value belongs to (0,1], it shows that the global space is positively correlated, and similar observed values have agglomeration effect, being more significant when values being closer to 1. When Moran’s I value is equal to be 0, it is shown that the space is not relevant and the attribute values are randomly distributed.

2.1.2 Local spatial autocorrelation test
The autocorrelation coefficient of the global space can explain the agglomeration effect of the space, but the discretization degree of the local space cannot be described, at the same time, it assumes the stability of the space, which cannot be based on the real space. Therefore, local spatial autocorrelation analysis is needed. In general, spatial heterogeneity is explained by the LISA (Local Indicators of Spatial Association) significance level and the Moran scatter plot. And The LISA index is usually analyzed by the local Moran’s I value, as shown in equation (2):

\[
Moran's I_l = \frac{(X_i - \bar{X}) \sum_{j=1}^{n} W_{ij} (X_j - \bar{X})}{S^2}
\]  

In the formula (2), \(S^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2\), the research areas involved are all local units. The Moran scatter plot mainly analyzes the instability of the space, and divided into four quadrants. The first quadrant represents high observed value units are around the high observed value areas. The third quadrant expresses low observed value units are around the low observed value areas. The two
quadrants show the positive dependence of the local space. Similarly, the second quadrant represents 
low observed value units are around the high observed value areas, and the fourth quadrant is the 
opposite. The two quadrants indicate the negative correlation of the local space, namely heterogeneity.

2.2 Spatial econometric analysis

2.2.1 Spatial Lag Model
Spatial Lag Model (SLM) mainly studies whether variables have spillover effects in a region, in other 
words, the influence of neighboring dependent variables on this region, as shown in equation (3):

\[ Y = \rho W Y + X \beta + \varepsilon \]  

In the formula (3), “\( Y \)” means the vector including explained variable with \( n \times 1 \) dimensions, “\( X \)” indicates the matrix, exogenous explanatory variables with \( n \times k \) dimensions included, “\( \rho \)” expresses 
the spatial hysteresis autoregressive coefficient, if it is significant, then the spatial dependency of lag 
term exists, “\( \beta \)” represents parameter vector with \( k \times 1 \) dimensions, “\( \varepsilon \)” means the random error term.

2.2.2 Spatial Error Model
Spatial Error Model (SEM) measures the influence of the error of the dependent variable which 
belongs to adjacent region on the observed value of the region, as shown in equation (4):

\[ Y = \lambda W \mu + \beta X + \varepsilon \]  

In the formula (4), \( \mu \sim N(0, \sigma^2 I) \), “\( \mu \)” represents the random error term of normal distribution, “\( \lambda \)” means the spatial autoregressive coefficient, if it is significant, then the spatial dependency of the error 
term exists.

3. Empirical analysis

3.1 data selection
This paper chooses the panel data of 31 provinces, cities and autonomous regions in Chinese mainland 
from 2010 to 2016, and makes an empirical analysis on the relationship between highway construction 
investment and regional economic growth. Except for individual special data, all the data required in 
this article are taken from the official website of “National Bureau of Statistics”, “China Statistical 
Yearbook”, “China Transportation Statistical Yearbook”, “China Education Statistical Yearbook” and 
relevant statistical yearbook of each province. Processing and analysis of data are done through the 
software of Geoda and Matlab.

3.2 Variables setting

3.2.1 Variable of regional economic growth selection
The variable of regional economic growth is commonly measured by GDP index, being divided into 
gross index and per capita index often. Considering the unsuitable for direct comparison of gross index, 
and the per capita index can exclude this adverse effect well. Therefore, this paper chooses per capita 
GDP as the explained variable to measure the difference of regional economic growth level. The ratio 
of regional GDP to total regional population expresses the per capita GDP index (unit: ten thousand 
yuan per person).

3.2.2 Variable of highway construction investment selection
Highway construction is largely controlled by Chinese government, not entirely following with the 
market economic standards, so highway construction investment variable cannot be measured in 
currency form, should be taken in kind form to measure [7]. Due to the differences on area of each 
region and the exclusion of idle highway resources, some literatures related tend to select the highway 
operating mileage density. Given the differences between highway operation and highway
construction, this paper chooses the highway network density (abbreviated as GLMD) index as a variable to reflect highway construction investment. As the explanatory variable in the model, highway network density index can not only show the results of highway construction, but also can reflect the transport capacity. The ratio of regional highway mileage to regional population expresses the highway network density index (unit: km per ten thousand people).

3.2.3 Control variables selection

According to the principles of economics, the main factors affecting the regional economic level are as follows: fixed-asset investment, abbreviated as FA (unit: ten thousand yuan), human capital, abbreviated as HR (unit: ten thousand people), labor force, abbreviated as LAB (unit: ten thousand people), household consumption level, abbreviated as SR (unit: ten thousand yuan), the ratio value of tertiary industry output to GDP, abbreviated as TI (unit: %), government expenditure, abbreviated as GOV (unit: %), trade openness degree, abbreviated as TRA (unit: %), etc.

Among them, the government expenditure is the result of government's policy intervention consciously, instead of market equilibrium. As for the factor of trade openness degree, there are obvious differences between the east and west in China, not considered as one of control variable in the paper. Taken into account of the operability of empirical research, the availability of data related and the strength of influencing factors, this paper selects the human capital, fixed-assets investment, labor force and the ratio value of tertiary industry output to GDP as the control variables ultimately. The variable of human capital estimates the level of education in a region, and measured by Education Indicator Act, in which includes average years of schooling index, literacy rate index, admission rate index and so on. Among them, literacy rate represents education level of population and sustainable development capacity in a country, and adult literacy rate could describe human capital’s contribution on regional economic growth accurately. Therefore, adult literacy rate which would be reckoned by illiteracy rate of people aged 15 years and above is selected for measuring human capital indicator. The number of employed persons at the end of the year in three industries represents labor force variable specifically. The fixed-asset investment variable is quantified as the actual investment amount of fixed assets in each year. In order to reduce the possibility that deviation of the analysis results exists owing to the data unit inconsistency. Logarithm is taken for all data of empirical research.

3.3 Empirical analysis of spatial statistics

3.3.1 Global spatial autocorrelation test

Taking the per capita GDP index as the explained variable, the data of which are taken from 31 provinces, cities and autonomous regions in Chinese mainland from 2010 to 2016, and the spatial correlation test is carried out by using Moran's I index with Geoda software to generate Moran's I scatter plot of the per capita GDP index of 31 provinces, cities and autonomous regions from 2010 to 2016, the specific test results in the year of 2010, 2012, 2014, 2016 are picked and shown in figure 1.

![Figure 1. Moran's I scatter plot of per capita GDP in the year of 2010, 2012, 2014, 2016](image)

In figure 1, the Moran's I values of per capita GDP are greater than zero in 2010, 2012, 2014, 2016, which indicates the explained variable of per capita GDP has spatial positive correlation, and the significance on positive correlation of spatial correlation is still to be further tested, then results of significance are shown in table 1.
Table 1. Moran's I value and Z-value for per capita GDP in 2010-2016

| Year | Moran's I | E(I) | Sd  | Z-value | P-value |
|------|-----------|------|-----|---------|---------|
| 2010 | 0.280     | -0.028 | 0.098 | 3.079  | 0.01    |
| 2011 | 0.277     | -0.028 | 0.095 | 3.157  | 0.01    |
| 2012 | 0.280     | -0.028 | 0.094 | 3.263  | 0.01    |
| 2013 | 0.281     | -0.028 | 0.091 | 3.502  | 0.01    |
| 2014 | 0.282     | -0.028 | 0.098 | 2.793  | 0.01    |
| 2015 | 0.286     | -0.028 | 0.092 | 3.257  | 0.01    |
| 2016 | 0.282     | -0.028 | 0.091 | 3.571  | 0.01    |

According to Table 1, the normal statistic Z-values of the global Moran's I during the period of 2010-2016 are greater than the critical value (2.56) at a significance level of 1%, which represents the significant level of Moran's I index test passes, therefore, there is a significant positive correlation on the explained variable of per capita GDP between various provinces and municipalities.

3.3.2 Local spatial autocorrelation test

The global spatial autocorrelation shows that the index of regional economic growth has spatial autocorrelation characteristics, but it cannot explain the status of spatial agglomeration and discretization between native and neighboring provinces, and the local spatial autocorrelation can make up for the deficiency. Based on the data in 2010 and 2016, the local spatial autocorrelation of regional economic growth index is tested by Geoda software. The results related are shown in Figure 2.

![Figure 2. LISA chart of per capita GDP of provinces and municipalities in 2010 and 2016](image)

According to Figure 2, the spatial agglomeration and discretization effects on regional economic growth index exist in 2010 and 2016. Particularly, in 2010, the provinces and municipalities of Beijing, Tianjin, Hebei, etc., being Beijing-Tianjin-Tangshan region, show the situation of high-high agglomeration, which indicates the region forms a good pull space pattern on economy with its surrounding provinces and the clear spatial dependencies appear. The economy base in this region is solid, and the network pulling economy has already been formed. Therefore, there is a realistic basis for the situation of high-high agglomeration. Guangdong, as the representative of the Pearl River Delta region and the southeast developed provinces, shows the situation of high-low agglomeration, indicating it experiences a rapid economic growth while the neighboring provinces are still in the low economic growth dilemma, being in a poor interaction and spatial heterogeneity status with its surrounding provinces. For a long time, Guangdong benefits from the reform and opening- up policy, coupled with the coastal advantage, its economic growth performance is doing well compared with the neighboring provinces, realistic basis for spatial heterogeneity exists. By 2016, the Beijing-Tianjin-Tangshan region continues to maintain the status of high-high agglomeration, The Yangtze River Delta region has a good economic interaction and the situation of high-high agglomeration is showed.

Except for the areas above, the characteristics of spatial dependence and heterogeneity of other regions are not significant in 2010 and 2016. From the view vertically, there is little changes of the economic agglomeration and discrete characteristics between the province and its neighboring provinces in 2010 and 2016, being at a stable level. Due to the influence of national economic policies, there is possibility that the regions of strong economic interaction occur. Looking horizontally,
regional differences in economic growth are distinct, the developed areas of eastern are more interactive in terms of economic growth, especially, the Beijing-Tianjin-Tangshan region, the Yangtze river delta area and the Pearl River Delta region. The synergies between the provinces in the western provinces and the neighboring provinces need to be further enhanced as well as the middle areas.

3.4 Empirical analysis of spatial measurement

According to the results of spatial statistical analysis, it is concluded that the explained variable has spatial dependence and heterogeneity. Then the spatial econometric model is chosen. In order to test the applicability of the spatial lag model or the spatial error model in this paper, the ordinary least squares regression, the residual Lagrangian Multiplier (LM) and Robust statistical test are used for carrying out verifying research. Anselin & Florax (1995) proposed the selection criteria of the spatial lag model and the spatial error model. In terms of LM-lag statistic and LM-err statistic, if the former is statistically significant, but the latter not being significant, or both are significant, but the former is more significant, then the spatial lag model can be selected as the applicable model; Otherwise, the spatial error model will be chosen as the applicable model. For that reason, MATLAB software is applied for achieving the estimation and test on ordinary model of panel data, and the specific results are shown in table 2.

| Variable     | Unfixed Effect | Time Fixed Effect | Space Fixed Effect | Time and Space Fixed Effect |
|--------------|----------------|-------------------|-------------------|-----------------------------|
| Ln RJGDP     | 10.823         |                   |                   |                             |
| Ln GLMD      | -0.401         | -0.416            | 0.559             | 0.167                       |
| Ln HR        | -0.141         | -0.070            | -0.036            | 0.208                       |
| Ln LAB       | 0.083          | 0.078             | -0.014            | -0.308                      |
| Ln FA        | 0.466          | 0.392             | 0.509             | 0.168                       |
| Ln TI        | 0.575          | 0.499             | 0.598             | 0.256                       |
| R Square     | 0.691          | 0.679             | 0.892             | 0.375                       |
| Log -L       | -3.240         | 4.192             | 235.582           | 336.470                     |
| LM-lag       | 3.721          | 0.054             | 0.020             | 0.887                       |
| R LM-lag     | 19.348         | 0.000             | 11.855            | 0.001                       |
| LM-err       | 0.896          | 0.344             | 6.741             | 0.000                       |
| R LM-err     | 16.523         | 0.000             | 18.576            | 0.000                       |

As can be seen from table 2 that, in the estimation and test results of the ordinary model’s panel data, compared with the model ineffective, the model’s R-square value (0.892) with the space fixed effect is greater than the model’s R-square value (0.691) without fixed effect, the R-square value (0.679) with time fixed effect and the R-square value (0.375) with space fixed effect, that is, the model of space fixed effect is with the best fit degree. And the log-likelihood value (235.582) of the model with space fixed effect is larger than the log-likelihood value (-3.240) of the model without fixed effect and the log-likelihood value (4.192) of the model with time fixed effect. Though the Log-likelihood value (235.582) with space fixed effect is less than the same value (336.470) with time and space fixed effect, the former’s value is large enough. Considering the model value (0.375) of fitting degree with space time fixed effect is low, so the model with space fixed effect is selected preliminarily in this paper. Further, in table 2, the model values of LM-lag, LM-err, Robust LM-lag and Robust LM-err with space fixed effect all pass the significance level test of 1%, and the LM-lag value (116.620) is greater than the LM-err value (86.265), meaning the former’s test statistic is more significant. Therefore, the SLM model with space fixed effect is chosen as the applicable model, and the model’s regression and test are carried out. Next, Matlab software is used for executing SLM.
model’s regression and test process, the specific results are shown in Table 3.

### Table 3. Estimation and test results for SLM model

| Variable | Estimation Coefficient | T Value | Direct Value | T Value | Indirect Value | T Value |
|----------|------------------------|---------|--------------|---------|----------------|---------|
| Ln GLMD  | 0.195                  | 1.987*  | 0.243        | 2.085*  | 0.130          | 1.995*  |
| Ln HR    | 0.103                  | 2.698** | 0.129        | 2.831** | 0.075          | 2.732** |
| Ln LAB   | 0.018                  | 0.201   | 0.015        | 0.139   | 0.005          | 0.127   |
| Ln FA    | 0.121                  | 3.918***| 0.152        | 4.305***| 0.117          | 4.834***|
| Ln TI    | 0.176                  | 2.230*  | 0.287        | 2.475*  | 0.129          | 1.973*  |
| R Square | 0.991                  |         |              |         |                |         |
| W*dep. var| 0.731                |         |              |         |                | 19.037  |
| Log Likelihood | 329.739        |         |              |         |                |         |

Note: "*", "**", "***" respectively represent the significance level of 10%, 5%, 1%.

From Table 3, except for the LAB variable, the variables of GLMD, HR, TI and FA all pass the significance level test, which shows the latter are the main influences for the regional economic growth. To be specific:

Firstly, the R-square value estimated by the SLM model is 0.996, fitting degree of which is relatively ideal. It is of great practical significance to describe the relationship between highway construction investment and regional economic growth.

Secondly, the variables, which include the factors of GLMD, HR, TI and FA, of estimation coefficient, direct effect coefficient and indirect effect coefficient are all positive, indicating these factors have positive effects on the regional economic growth of the province and neighboring provinces, and there is no negative spatial spillover effect. The coefficient of GLMD factor is estimated by 0.195, which represents the change of 1% on GLMD factor will cause the change of 19.5% on regional economic growth in the province positively. Similarly, 1% change on FA factor input will bring about the economic growth by 12.1% positively, 1% change on TI factor input will result in the economic growth by 17.6% positively.

Thirdly, the direct effect of GLMD input factor is greater than the direct effect of HR, FA and TI, and it has a great influence on the spatial pattern of economic growth in various provinces and regions. For the local government, GLMD factor input should be given priority for expanding the input-output ratio. In addition, the input factors’ direct effect and indirect effect of GLMD, TI, FA, HR decline in turn, compared with the input factors of HR and LAB, the input factors of GLMD, FA and TI show stronger pulling effect on the regional economic growth. For the input factors of GLMD, FA and TI, their influence on economy in the province and adjacent provinces are more direct and fast, and due to the slow conduction, HR input factor influences slowly on economy promotion.

Finally, the direct effect of each input factor is less than its indirect effect, which shows that when a unit of factor input is added in the province, greater economic benefits will be obtained in the province than the neighboring province, being in line with the reality of economic development pattern.

### 4. Conclusion

There is significant spatial autocorrelation in the per capita GDP of various provinces and municipalities. Strong spatial dependencies exist between some provinces and cities, especially, the economic growth in the eastern area shows obvious effects on interaction, such as the provinces of Hebei, Beijing, Tianjin, Shanghai, Zhejiang and other areas, and appears the trend stability. By contrast, the pull effects of economic growth in the province of Guangdong for neighboring provinces are weak.

According to the regression results on direct effect of SLM model, these factors which include highway construction, human capital, labor force, fixed-asset investment and the ratio value of tertiary industry output to GDP perform obvious positive effect on regional economic growth. Except for the factor of labor force, the promotion effects on the others of regional economic growth are significant,
the factor of highway construction being prominent especially. Simultaneously, in the study of elements, the regression results on indirect effect of SLM model reveal that, compared with other factors, the input factor of highway construction brings the greatest effect of the economic benefits on its neighboring provinces. The reason for this phenomenon is that the existence of social benefits on highway connectivity, which is beneficial to the economic trade and communication between native and neighboring provinces, even neighboring provinces will achieve a lot economic benefits with smaller economic costs for regional economic growth further. From another aspect, the empirical analysis in this paper confirms the macro management philosophy of government: "Building the highway is the first step to become rich".

References
[1] Ye C Y and Wang D X. Traffic infrastructure, transportation and regional economic growth——Based on spatial panel model of provincial data [J]. Industry Economics Research, 2013 (02): 45-46.
[2] Zheng G J. Traffic infrastructure, spatial structure adjustment and regional economic coordination[D]. Zhejiang University, 2014: 164-169.
[3] Zhang X L. Does China's transport infrastructure promote regional economic growth? —— Review on the spatial spillover effects of transport infrastructure [J]. Chinese Journal of Social Sciences, 2012 (03): 69-70.
[4] Yu J X. Spatial Difference and Correlation Analysis of Highway Transportation and Economic Development in China[J]. Journal of Transportation Systems Engineering and Information, 2015 (01): 11-16.
[5] Chu C L, Zhang Z H and Xie N. Study on spatial spillover effect of expressway construction on regional comprehensive development [J]. Journal of China Highway, 2015 (07): 93-100.
[6] Li Z Q, Ou G L and Mao G Y. Study on spatial correlation of highway traffic infrastructure and regional economic development [J]. Journal of Yunnan University of Finance and Economics, 2016 (01): 50-60.
[7] Hu J J and Sun J W. Spatial measurement——model, method and trend[J]. World Economic, 2014 (06): 111-115.
[8] Guo X L. Spatial spillover effect of transport infrastructure on regional economic growth in China——Based on empirical research of provincial data [D]. Beijing Jiaotong University, 2014: 45-46.
[9] Moreno R, Lopezbazo E. Returns to Local and Transport Infrastructure under Regional Spillovers [J]. International Regional Science Review, 2007 (01): 60-70.
[10] Elhorst J P. Specification and Estimation of Spatial Panel Data Models[J]. International Regional Science Review, 2003 (03): 250-264.