IMPACTS OF PUBLIC TRANSPORTATION ON THE URBAN RURAL INCOME GAP IN CHINA

by

Zhuo Zhang

Submitted in partial fulfilment of the requirements
for the degree of Master of Arts

at

Dalhousie University
Halifax, Nova Scotia
August 2014

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The urban rural gap is one of the most changing economic problems that confronting China today. I study this problem by focusing on the impacts of public transportation density on the urban rural income gap using provincial panel data from the last 20 years. I conduct the spatial econometric model in this study to deal with the transportation spillover effects. I use GDP, the degree of openness, and average years in school as my control variables in my estimation. The Result shows increasing transportation density negatively contributes to the urban rural income gap.

LIST OF ABBREVIATIONS USED

RMB	Renminbi
OLS	Ordinary Least Squares
SAC	Spatial Auto Correlation

ACKNOWLEDGEMENTS

Dr. Mutlu Yuksel: without your encouragement and criticism regarding this thesis, I could not finish this project.

Professor Andrea Giusto and professor Mevlude Akbulut-Yuksel, I am very grateful for your valuable suggestions and corrections.

Kerrianne Ryan at the Dalhousie Writing Centre: you give me valuable suggestions on writing and structure and future research topics.

My father and mother: without your financial support and mental support, I would not have the opportunity to study at this amazing university – Dalhousie.

Chapter 1. Introduction

It has been over three decades since China opened its economy to the global market and started economic reform in 1978. Over this period, China has produced high economic growth and over 500 million people have risen above the poverty level. Although these gains are quite impressive, they have been accompanied by a rise in income inequality. A high level of inequality creates an unfavorable environment for economic growth and is bad for political stability. Although China is known as the largest socialist country, and set a goal of becoming a "Socialist New Country" without inequality, it has instead become one of the countries with the largest income inequality. In fact, China shows significant evidence that the income inequality surpass the United States – the largest capitalist economy in the world (Milanovic, 2005). Comparing income inequality levels to those of other middle-income countries like Thailand and Malaysia, reveals that this problem in China is much worse (Adams, Arvil Van. 2009). This inequality is reflected in the increase of the Gini coefficient from 0.28 in the early stages of economic reform to 0.42 at present (World Bank, 2009).

Decomposition of China's income inequality by Sicular, Yue and Gustafesson (2007) reveals that the urban rural income gap is the main contributor to overall inequality in China. In 2002, the urban rural income gap contributes 45 percent to the overall income gap; and in 2007, the proportion increases to 51 percent. (Li et al. 2013). The urban rural income gap itself has also grown, so that the ratio of average disposable income in urban areas over net income for rural areas rose from 1.8 in 1983 to 3.1 in 2012 (Calculated from data in Statistic China). Thus, the extensive contribution of increasing urban rural income gap to income inequality suggests that it is now an emerging task to reduce the urban rural income gap in China.

Due to the productivity increase in the agriculture sector, a large work force has now become available. These available rural workers tend to move to the manufacturing and service sectors, which are located in the urban areas. The surplus of 100 million internal migrant workers has increased the labour supply and has brought down wage levels. This flow of workers has boosted the industrial economy. Many move to urban areas for work, but require access to their families in rural areas. Thus, transportation has played a very important role in this labor migration wave.

Transportation not only has played an essential role in the labor migration wave, it has been an important channel to boost the economy. In response to the Asian financial crisis happened in 1998, the Chinese government conducted a strong fiscal policy- increasing public spending on railway and highway construction. Up to 2008, China's total length of rail way ranked number 3 in the world and total length of highway ranked number 2 in the world (Huang, Yu and Wei, 2013). In response to the 2008 financial crises, Chinese government came up with 3 trillion RMB public projects. Public transportation accounts for the largest proportions of these projects. If the increased transportation infrastructure made income inequality more serious, then government should think twice when they make similar decisions. However, if transportation infrastructure does the opposite, then building more roads is good news for reducing the urban rural income gap.

Developments of transportation infrastructure can play important role in reducing the urban rural income gap. There are mainly four channels through which transportation affects income distribution. First, public transportation provides people living in rural areas with access to opportunities in urban areas. They can travel by trains or buses to work in cities, therefore increasing their wages. However, the outcome of this access on income gap is ambiguous. If increased labor supply in cities increases the income of people in urban areas more than the wage increase of rural workers, then this will enlarge the income gap. Secondly, transportation can stimulate the process of production specialization, increasing productivity in that area and increasing the wages of workers from rural areas. Thirdly, people from rural areas can benefit from transportation by increasing their human capital. Transportation can increase their probability of access to good education, thereby increasing their chance of getting better jobs in the future. Last but not least, transportation infrastructure can transport farm products to urban areas. The ability to transport products means farmers have a larger market than before. In supermarkets, people can buy farm products from all over the country. Thus, I expect that transportation density will have a positive impact on the urban rural income gap.

I examined the impacts of transportation density on the urban rural income gap using the method of spatial econometrics. I use provincial panel data from 1993 to 2012 obtained from China Statistic Yearbook. The dependent variable, the urban rural income gap, is measured using the ratio of average disposable income in urban areas over net income for rural areas. This is similar to Huang, Yu and Wei(2013)'s method to calculate the urban rural income gap. The independent variable, transportation density, is the total length of railway and highway in one province divided by the total area of that province. Both direct effect and indirect effect are measured; direct effect refers to the impacts of transportation density on its own province's urban rural income gap. Indirect effect refers to how an increase in a certain province's transportation density can influence the urban rural income gap in other provinces. I control for education level in each province, degree of openness, real GDP per capita and square of real GDP in this study. In addition to the spatial model, I also include OLS results for comparison. Several key findings emerge from my analysis. The main finding is that building an intensive transportation network help to reduce income inequality because it causes the borders of urban and rural to become blurred. The analysis also provides evidence that railways play a more significant role in reducing the urban rural income gap than highways. Additionally, I find the spillover effects, known as the indirect effects, are more significant than the direct effects. As for the implications of the findings, it is highly recommended that more railways should be built. For future relevant research questions, I suggest studying of the impacts of transportation density on coastal-hinterland income gap at the end of this paper.

In this paper, I contribute to the literature on transportation spillover effects on the urban rural income gap in three ways. First, I conduct spatial econometric models to deal with the spillover effects. Because of the spatial analysis, the provincial panel data in my study is no longer independent, so all provinces are geographically correlated with each other. This makes a lot of sense because, in reality, transportation is like a circulatory system which binds the economy as a whole unit. Second, although there are a lot of papers studying the urban rural income gap in China, as well as how transportation development can influence economic growth, few studies set their eyes on how transportation development can influence the urban rural income gap. Third, I updated my data set to the year 2012, which is the latest date available on Statistic China.

Chapter 2. Background and Literature Review

In this section I provide the background of China's registration system, which makes the urban rural income gap stand out from other countries, a quick overview of recent literature on the impacts of transportation development on inequality, several other key factors that may influence the urban rural income gap, and the uniqueness of this study from previous ones.

Before I go into detail, I need to first address the unique aspects of the urban rural income gap that make China's situation different from other countries – China's household registration or hukou system. The aim of adopting hukou system is to control domestic population movements. The hukou system acts as an internal passport system that hinders people from rural areas becoming official urban residents. As a consequence, short term and temporary consist most part of urban migration. (Sicular, 2013). Because of this special situation in China, a fast and convenient transportation network is needed for the migration wave in China. If the migration wave was permanent and once rural workers moved to cities they could be counted as urban residents, then they probably would take the trouble of taking trains and transferring several times to move into cities. Ishtiaque and Ullah (2013) explore the push and pull factors that influence rural-urban migration in Bangladesh. Transportation development is not a crucial factor that plays a role in the rural-urban migration (Ishtiaque and Ullah, 2013). However, because rural worker's urban residence is not permanent in China, they need a convenient transportation network to enable them to move to urban areas to work while still being able to visit their families in rural areas.

The impacts of the hukou system on the urban rural income gap in China makes its study unique from an economic perspective. Literature from other countries is more focused on the study of the effects of infrastructure development on overall income inequality instead of the urban rural income gap. Calderón and Servén (2004) find that infrastructure quantity and quality have robust negative and significant impacts on the Gini coefficient. The results are consistent with the hypothesis that infrastructure development strengthens the ability of people living in poverty to access additional productive opportunities (Calderón and Servén, 2004).

Using transportation as a way to reduce poverty seems to have greater impacts on poor rural areas than the urban areas. According to Gannon and Liu (1997), there are mainly four reasons why transportation can reduce the poverty in rural areas more effectively in the urban areas. First of all, rural poor are often geographically isolated and they are more homogenous group than the urban poor, so targeting can be relatively effective. Second, works like road maintenance or road building are very labor intensive, and such works may provide rural people with opportunities to earn extra money aside from framing. Third, because there are few opportunities in poor areas, rural people will have more incentive for participation, which is often key to success for direct interventions. Overall, transportations improvement is a very good method to alleviate rural poverty (Gannon and Liu, 1997).

To decide which factors other than transportation development should be included as control variables, I explore several previous studies on the contributors to urban rural income gap. Those include economic development, human capital and GDP (Kuznets, 1995; Baldi, 2013; and Anderson, Huang and Ianchovichina, 2004).

In addition to transportation, economic development can influence the urban rural income gap. Kuznets (1955) studies this field by investigating the distributional consequences of different stages of economic development. He finds that with the development of economy of a certain area, income inequality in that area will first rise, then stay at a stable level, follows by a decline. Barro (2000) finds evidence from a broad panel of countries shows that the turning point is $2000. That is to say: GDP growth tends to increase alongside inequality if GDP per capita is less than $2000 (1985 U.S. dollars), but above $2000, inequality decreases as GDP increases. Thus, Barro (2000) show that there appears to be a global trend based on GDP, rather than specific National circumstances.

The urban rural income gap can also be influenced by human capital. Baldi (2013) finds that better educational opportunities can help economic development and help to reduce income inequality. In theory, development of education institution can benefit the poor (people in rural area in our case) or the rich (city residents). For example, if development of educational institution development can help people, who do not have access to education institution before, have opportunities to gain better education now. This will reduce income gap. However, if only those who have already got access to education institution,   In my model, I will check which one has greater effects.

Trade is another factor that can have impacts on income inequality. Anderson, Huang and Ianchovichina (2004) talk about the impacts of China's WTO accession on rural-urban income inequality. They find that income for famers from nonfarm work will be greater than the income increase for urban people. Because wages for unskilled workers in rural non-farm activities will rise, they believe rural non-farm poverty will fall. I think development of public transport will help farmers to transport their products to cities. Additionally, by joining WTO, demand for both goods and will increase. Better public transportation will help people in rural areas to get to places where laborers are needed.

Taking all the factors I discussed previously which influence the urban rural income gap into consideration, Huang, Yu and Wei (2013) try to answer the question – can transportation infrastructure decline urban rural income gap. They use provincial panel data from 1991 to 2007 in China. They use a spatial lag model and spatial error model to capture the spillover effects of public transportation. They find transportation infrastructure declines the urban rural income gap and the effect of the highway infrastructure on the urban rural income gap is bigger than that of railway. Although I am studying the same topic as them and I also find transportation contribute negatively to the urban rural income gap in China. I have made some improvements on Huang, Yu and Wei's(2013). First, I use the latest data up to 2012, which updates information in Huang's study which was done in 2007. Second, I conduct a different spatial econometric method – spatial autocorrelation model. This model is better than spatial lag model and spatial error model in this case of study. Third, the railway plays a more important role in reducing the urban rural income gap based on my study, which is quite different from Huang's results. Last but not least, I confirmed the existence of inverted U shape Kuznets relationship in China, which Huang Yu and Wei(2013) do not confirm.

This paper examines the impacts of transportation density on the urban rural income gap. Using spatial econometric model as well as OLS, I analyse the spillover effects of transportation network, taking on a completely different approach than previous work done by Huang, Yu and Wei(2013). This research combines studies on the key factors which influence income distribution (the urban rural income gap in my case) to select the control variables.

Chapter 3. Models and Methodology

3.1 Spatial Specification

The starting point for my discussion is a simple OLS regression. The scatter plot of income gap against transportation density reveals a negative relationship between them can be seen. See figure 2. The OLS regression is:

$$Incomegap_{\mathrm{it}}= \alpha_0+\alpha_1 trans^{e}_{\mathrm{it}}+\alpha_2 Z_{it}+\varepsilon_{it} \tag{3.1}$$

The dependent variable is the rural-urban income gap. The independent variable $trans^{e}_{it}$ represents effective public transportation stock. I use railway density and highway density to represent $trans_{it}$. Z_{it} is the set of control variables, including average years in school, degree of openness, real GDP and square of real GDP. All the variables are in provincial level. The coefficient for $trans^{e}_{it}$ (α_1) is the value of interest in this study. If α_1 is negative and significant, then it indicates that increasing public transportation infrastructure will decrease the income gap between urban and rural areas.

Holtz-Eakin and Schwartz (1995) propose an approach to “the effective stock of transportation” expressed as $trans^{e}_{\mathrm{it}}$, which includes spatial spillovers. To deal with spatial spillover effect, I use spatial econometrics. It incorporates spatial interaction effects among geographical units. Spatial econometrics requires the creation of a spatial weight matrix to describe the spatial arrangement of the geographical provinces in China. The effective stock of state railway and highway capitals differs from the physical stock of highways and railways within that province’s borders. That is to say, access to the railways and highways in other provinces also contributes to the effective stock of public transportation in a given province, which will lead to the effective stock exceeding the physical stock within the province’s borders (Holtz-Eakin and Scheartz’s, 1995). Holtz-Eakin and Scheartz (1995) define the effective capital stock of a province(province i) as the sum of physical transportation stock in that province and the effective transportation stock in its neighboring province (province j), expressed by equation:

$$trans^{e}_{it}=trans_{it}+\rho trans^{e}_{jt} \tag{3.2}$$

$transt_{it}$ is the physical stock of transportation in province i. If there are no interprovincial spillovers, ρ=0 and the effective and actual transportation stock measures will coincide ($trans^e_i = trans_i$). If 0<ρ<1, it means that part of neighboring effect transfers to province i. If ρ=1, it means that all of the neighboring province's effective capital spills over to the effective transportation stock of province i.

Equation (3.2) assumes that only one province neighbors province i. In reality, it is safe to consider that i has more than one neighbor. So equation 3.2 is expanded to:

$$trans^e_{it}=trans_{it}+\rho\sum_{j=1}^{N} trans_{jt}*w_{ij} \tag{3.3}$$

N is the total number of provinces in the data set. w_{ij} is the weight assigned to province j. I create a spatial weight matrix that is a 31×31 "neighbors" matrix containing weights for each state's transportation capital. There are two types of spatial weight matrices (W): 0-1 spatial matrices and inverse-distance spatial weight matrices. If w is a 0-1 spatial matrix, then

$$w_{ij}=\begin{cases}1 & if\ i\ and\ j\ are\ neighbors\\ 0 & if\ i\ and\ j\ are\ not\ neigobors\end{cases} \tag{3.4}$$

In an inverse-distance spatial matrix:

$$w_{ij}=\begin{cases}1/d_{ij}, & if\ i\neq j\\ 0\ , & if\ i=j\end{cases} \tag{3.5}$$

d_{ij} is the distance between the capital cities of province i and j. Normalizing the spatial weight matric requires that every row in the matrix adds up to one. According to Holtz-Eakin and Scheartz (1995), the physical transportation stock in province i not only contributes to its own province's effective transportation stock, but also contributes to its neighboring provinces through a rate ρ. In addition, it contributes to the subsequent surrounding ring of provinces at a rate ρ^2 and so on to ρ^3 etc. If ρ is high, closer to 1, it means slow decay of the spillover effect, and vice versa (Holtz-Eakin and Scheartz's, 1995).

Normalizing the matrix results in the normalized spatial weight matrix—W. T denotes the ($N\times 1$) vector of transportation stock for the provinces; W denotes the normalized spatial weight matrix. Equation (3.3) can now be rewritten in matrix form:

$$T^e=T+\rho WT^e \tag{3.6}$$

Solving equation for T^e yields:

$$T^e=(I-\rho W)^{-1}T \tag{3.7}$$

Then substituting (3.7) to (3.1) I get:

$$\text{Income gap}= \alpha_0+\alpha_1 (I-\rho W)^{-1}T +\alpha_2 Z+\varepsilon \tag{3.8}$$

This takes care of the spillover effect arising from transportation. A new problem has been created in (3.8): the coefficient for ***T*** is non-linear. To deal with the new problem, Holtz-Eakin and Scheartz (1995) come up with a solution to multiply (I- ρW) by both the RHS and LHS which yields:

$$(I-\rho W)\text{Income gap}= \beta_0+\beta_1 T +\beta_2 Z+e \tag{3.9}$$

or,

$$\text{Incomegap}= \beta_0+\rho W^*\text{Income gap}+\beta_1 T + \beta_2 Z + e, \quad \text{where} \quad e=\lambda We+ \varepsilon \tag{3.10}$$

Equation (3.10) is a Spatial Auto Correlation model (SAC). W*Incomgap is a spatial lag dependent variable. λ is the coefficient which allows the spatial structure appears in the error term. If $\lambda=0$, then it is a spatial lag model. If $\rho=0$, then it is a spatial error model. Figure 1 displays all the major spatial models in spatial econometrics. I use maximum likelihood method to estimate all the coefficients.

3.2 Direct and Spillover Effect of Transportation

Direct effect of transportation refers to the impact of changes in physical transportation in province i on the urban-rural income gap in that province. Mathematically speaking, direct effect is the partial derivative of the income gap in province i with respect to the physical transportation stock in province i:

$$\text{Direct Effect} = \frac{\partial incomegap_{it}}{\partial trans_{it}} \tag{3.11}$$

The indirect effect, also known as the spillover effect of transportation is the impact of changes in physical transportation in province i on all other provinces' income gap. In mathematics, it is the sum of partial derivatives of the income gap in other provinces with respects to the physical transportation stock in province i:

$$\text{Indirect Effect}=\sum_{j=1}^{N-1} \frac{\partial incomegap_{jt}}{\partial trans_{it}}, \quad \text{where } j \neq i \tag{3.12}$$

I will use equation (3.11) and (3.12) to calculate the direct effects and spillover effects of public transportation on the urban rural income gap in China in the next Chapter.

Chapter 4. Data and Descriptive Statistic

I use annual provincial panel data from 1993 to 2012 in my analysis. I acquire data from the China Statistics Yearbook.[1] Thirty one provinces are included in my data set. I do not include Hong Kong, Macau and Tai Wan because of absent data.

The dependent variable is urban-rural income gap. In this study, dependent variable is measured by per capita annual disposable income of urban households dividing per capita annual net income of rural households. As this ratio gets larger, it indicates a greater rural-urban income gap in China. The main contents of the survey include the basic condition of rural households, housing conditions, income, consumption expenditure, consumption of major consumer goods and the quantity of durable consumer goods owned. Figure 3 shows the overall income gap in the 20 most recent years throughout the whole country. Figure 4 shows that during the ten most recent years, the income gap in provinces which are located in the west of China is the greatest whereas the east of China has the lowest income gap. This is quite interesting because the eastern part of China is considered to be the most developed part. There are 11 provinces in the east, including the capital city, Beijing and the financial center, Shang Hai. Compared with the east coast, the west of China is recognized as the most undeveloped part.

The variable of interest is public transportation infrastructure density. I use the sum of railway density and highway density to represent public transportation infrastructure density. I calculate the railway (highway) density by dividing the total length of railways (highways) in operation in one province by the total area of that province. The unit of railway density and highway density is kilometers per square kilometer. The reason for choosing railway and highway as the main explanatory variables is that, together, railways and highways transported to 88.6% of goods and 97.8% of passengers in China during 1991-2007 (Huang, Yu and Wei, 2012). The overall length of both highways and railways increases during the last 20 years. In particular, in the year 2005 there is a huge jump in the

[1] Data are available from the online website: http://www.stats.gov.cn/enGliSH/.

total length of highways. To ensure accuracy of the coefficient of transportation density, I control for degree of openness, human capital and GDP.

At present, there is not a common indicator which reflects the level of openness of any province. To overcome this issue, similar to Huang Yu and Wei(2013)'s paper, I use the ratio of total value of imports and exports of a province to the total GDP in that province as an indicator to represent degree of openness. Hu (2000) in his paper finds that the income gap between the coastal area and the hinterland in China may be caused by the increasing rural- to-urban labor mobility and the improvements of trade conditions. Coast areas have great advantages for international trade because water transportation is much cheaper than railway transportation. Thus, coastal cities were picked to be the initial locations for industrial agglomeration. Moreover, increasing return to scale further confirms their leading positions. Hu's study (2000) shows the labor supply for the industrial agglomeration in coastal areas mainly comes from intraprovincial movement instead of interprovincial movement. The location disadvantages of the hinterland bring less trade to the hinterland, which became a key reason for the enlarging income gap between coastal areas and hinterland (Dapeng Hu, 2000). Although my research interest is not the same as Hu's, trade is still considered to be a variable which has influence on the rural-urban income gap within one province.

Human capital is another confounding factor which affects the urban rural income gap. Human capital can influence income distribution through two channels: the composition effect and the wage compression effect. The first refers to how increased education level can first enlarge income inequality and then decrease it. The second refers to the fact that as more people get educated, the return on education drops, so inequality decreases. I use average years in school by province to represent human capital in each province. The elementary school accounts for 6 years, the middle school diploma accounts for 9 years, and high school diploma accounts for 12 years. People with university degrees or higher are considered to have 16 years in school.

I also include Real GDP per capita in my set of control variables. Hu (2000) provides the following summary about the impact of GDP on income inequality, at early stage of development, income differentials first increase, then stay stable for some time and then diminish in developed stage. This is called an inverted U-curve (the Kuznets curve). I also include square of real GDP per capita to capture the curvature and determine whether this phenomenon happens in China and whether some provinces have reached the decreasing horizon. I choose 1978 as the base year to calculate the real GDP per capita. The measurement unit is RMB.

Table 1 shows the descriptive statistics for income gap, transportation density and the main control variables used in the estimation.

Chapter 5. Estimation Results

The results are summarized in Table 2. Column (1) is the OLS results using fixed effect without considering the neighborhood's transportation spillovers. Column (2) adds neighbor provinces' transportation density into the regression. Column (3) and Column (4) use spatial autocorrelation models. The only difference between (3) and (4) is their spatial weight matrices, which are different: column (3) use 0-1 weight matrix whereas column (4) uses inverse-distance matrix.

The LR test determines whether spatial econometrics or OLS is a more appropriate model. The LR test is a right tail test that yields the joint probability that both ρ and λ are equal to zero. Under the null hypothesis, OLS is consistent and efficient; under the alternative hypothesis, OLS is inconsistent and SAC is consistent. The LR test statistic for the current data is 982.0609 and the critical value is Chi-square with a degree of freedom of 2 and a P-value of 0.000. This result is consistent with spatial econometric model.

Several spatial econometric models exist, and the appropriate model selection requires examination of spillover effects in the spatial lagged variable and the error term. To check the significance of rho and lambda separately, The LM lag (robust) test tests whether the spatial lagged dependent variable has spatial autocorrelation. Under Ho: Spatial lagged dependent variable (W*Incomgap) has no spatial autocorrelation; Under Ha: Spatial lagged dependent Variable has spatial autocorrelation. The p-value for the LM lag test is 0.000, so using a 0-1 weight matrix, results in rejection of the null hypotheses of no spatial lag autocorrelation. However, if using inverse distance matrix, the p-value for the LM lag test is 0.984, which is very large. For reasons of simplicity, I am now only focusing on the 0-1 weight matrix. Problems with the invers-distance matrix in my model are discussed later on.

Another spatial panel autocorrelation test is the LM error (robust) test. This test tests whether the error term has spatial autocorrelation. Under Ho: error has no spatial autocorrelation; under Ha: error has spatial auto correlation. From table 2, it is clear that both columns (3) and (4) give low p-values for LM error test. So the null hypotheses should be rejected.

All three tests suggest the use of the SAC spatial econometric model instead of OLS due to the inconsistency of OLS results and spatial spillover in both the spatial lagged dependent variable and the error term.

Column (3) in Table 2 illustrates that coefficients for all dependent variables are significant under 1% significant level. Thus, increasing public transportation density can reduce the urban-rural income gap.

Unlike one-dimensional autocorrelation, spatial correlation is multi-dimensional. To measure spatial autocorrelation I use Moran's I which was developed by Patrick Alfred Pierce Moran. The range for Moran's I is between -1 to 1. Negative values indicate negative spatial autocorrelation which indicates dispersion of the geographic units' arrangement. In the extreme situation where Moran's I is -1, it means the income gap is perfectly dispersed; if Moran's I is +1, it means perfect correlation of the income gap. Zero indicate a random spatial pattern.

> *Cited from Wikipedia, Moran's I is defined as ：“*
>
> $$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$
>
> *where N is the number of spatial units indexed by i and j; X is the variable of interest; $\bar{X}$ is the mean of X; and w_{ij} is an element of a matrix of spatial weights (Wikipedia,2013).”*

I calculate Moran's I for the dependent variable using both 0-1 spatial weight matrix and inverse-distance matrix. Both methods give positive significant Moran's I values. Table 3 provides the Moran's I values for the dependent variable for the past 20 years. All values for Moran's I are around 0.2 (Table 3). From 1998 to 2005, Moran's I increases gradually and has stayed at a high level since 2005; this may be caused by the economic integration. The income gap tends to cluster. One of the reasons might be that industry tends to cluster together. For example, provinces located in the south east of China are famous for their light industry, whereas heavy industry tends to cluster in the north east of China. Figure 5 gives a graphical analysis of how individual units cluster from a dispersed condition. Development of transportation infrastructure may help to stimulate economic integration by increasing the mobility of labor so as to increase the labor supply in urban areas. As a matter of fact, interprovincial labor mobility is very frequent in China (Hu, 2000).

Figure 6 shows that when the distance bands of a unit increase, the spatial correlation of the dependent variable decreases. This relationship makes sense because the impacts of a change in one place play little role in another place that is far away from it. The difference between the Moran's I correlogram in 2012 and that in 1993 is not substantial.

The marginal effects is displayed in table 4. By looking at the marginal effects, it is apparent that if transportation density increases by 1 standard deviation in province i, the income gap in province i will be reduced by -2.1031percentage points, and the income gap in all other provinces combined will be reduced by -3.18 percentage points in total. Therefore, the total effect is -5.29 percentage points.

Alternatively, the results in table 5 can be interpreted as a 1% increase in transportation density in province i corresponding with a -0.88% change in income gap in that province and a 1.33% decrease in income gap in other provinces. The coefficients for human capital and openness are negative, meaning that an increase in the education level and an increase in the volume of trade will decrease the income gap. Columns (3) and (4) have negative significant coefficients for square of real GDP, consistent with the inverted U shape Kuznets curve.

Next, I separate the transportation density into railway and highway density to check which one has greater impacts on the urban-rural income gap. The results are shown in table 6. The columns correspond to the alternative measures of transportation density introduced in table 2. Both columns (3) and (4) indicate that railways play a more important role in reducing the income gap. Because the coefficient for railway density is much larger than that for highway density. Additionally, column (4) suggests the highway density is not significant when we use inverse distance weight matrix. After separating transportation density into highway density and railway density, adjusted R-square improved for all the models except for the model of simple OLS without the correcting neighborhood effect (column 1).

Among all the models, spatial panel model 0-1 spatial weight matrix gives the lowest AIC and highest adjusted R-square. So column (3) in table 6 appears to be the best model.

I now briefly discuss why using the inverse distance weight matrix is not appropriate in this case. The Spatial Auto Correlation model deals with the problem that spatial correlation happens within the independent variable and this correlation decays as the neighbor gets farther away. The spatial inverse distance matrix, by construction, has already taken care of all the neighborhood effects, direct neighbors, and subsequent neighbors, and so on so forth. Using SAC along with the inverse distance weight matrix will duplicate the correction for these factors, this redundancy will provide false results. Moreover, the ρ value exceeds 1 which is not likely to happen.

Table 7 shows the direct and indirect marginal effect for each variable. The interpretation is as before when we interpreted table 5. One standard deviation increase in railway density in province i will decrease the income gap in province i by 2.79 percentage points and reduce 3.97 percentage points of income gap in all other provinces in total; one standard deviation increase in highway density in province i will decrease the income gap in province i by 1.56 percentage points and reduce 2.21 percentage points of income gap in all other provinces in total.

Table 8 shows the elasticity effects, the interpretation is similar to table 6: a 1% increase in railway density will decrease the income gap in that province by 1.01%, and a 1.44% decrease in all other provinces; a 1% increase in highway density will decrease the income gap in that province by 0.64%, and a 0.91% decrease in all other provinces.

Chapter 6. Conclusion

The increasing urban rural income gap is one of the most challenging economic problems confronting China today. I explored this problem by focusing on transportation, as well as other major factors associated with the widening urban rural income gap in China using provincial panel data from 1993 – 2012. Spatial spillover effects of transportation development were accounted for by conducting the spatial econometric method.

Some interesting results derived from my theoretical and empirical framework. My analysis suggests developing public transportation infrastructure by increasing transportation density will help to ease the urban rural income gap. In this process, railways play a more important role than highways in decreasing the urban rural income gap. Spillover effects of transportation density are greater than direct effects in helping to ease the urban rural income gap and I also find the inverse U shape Kuznets curve in China within past twenty years with the turning point of 14766 RMB. Other factors are also worth noticing in reducing the urban rural income gap. While improving education level will help to decrease urban rural income gap in China as it is significant under 1% significant level, the magnitude is small. More trade will also decrease income gap, and both the significance level and magnitude suggest that trade is important in reducing the income gap.

Further research about this topic can be done by separating China into several economic zones, east of China, middle China, and west of China, and examining how different transportation can influence the urban rural income gap in those three areas.

References

Adams, Arvil Van. 2009. Economic growth, spatial income inequality, and the role of education and training in China. South Asia : human development unit discussion paper series ; no. 29. Washington, DC: World Bank.

Anderson, Kym and Huang, Jikun and Ianchovichina, Elena, Will China's WTO Accession Worsen Rural Poverty? (January 2004). CEPR Discussion Paper No. 4196.

Asif Ishtiaque, Md. Sofi Ullah (2013). The Influence of factors of migration on the migration status of rural-urban migrants in Dhaka, Bangladesh, Journal of Studies and Research in Human Geography.

Branko Milanovic (2005), Half a World: Regional inequality in five great federations, World Bank Policy Research Working Paper 3699.

Colin A. Gannon and Zhi Liu (1997). Poverty and Transport. Discussion Paper, World Bank.

Dapeng Hu. (2000) Trade, rural–urban migration, and regional income disparity in developing countries: a spatial general equilibrium model inspired by the case of China. Regional Science and Urban Economics, volume 32, Issue 3.

Douglas Holtz-Eakin and Amy Ellen Schwartz. (1995). Spatial productivity Spillovers From Public Infrastructure: Evidence From State Highways. NBER working paper No.5004

Elhorst, J. P. (2013). *Spatial econometrics: From cross-sectional data to spatial panels*. Page 9

Guido Baldi(2013). Physical and Human Capital Accumulation And The Evolution of Income And Inequality. Journal of Economic Development

Kuznets, Simon. (1955). Economic Growth and Income Inequality. American Economic Review 45 (March): 1–28.

Moran, P. A. P. (1950). Notes on Continuous Stochastic Phenomena. *Biometrika* 37 (1): 17-23.

Robert J. Barro(2000). Inequality and Growth in a Panel of Countries, Journal of Economic Growth

Terry Sicular, Yue Ximing and Bjo¨rn Gustafsson, 2007, The urban-rural income gap and inequality in China, Review of Income and Wealth, Series 53, Number 1.

Yan Li and Maria N. DaCosta (2013) Transportation and income inequality in China: 1978–2007, Transportation Research Part A: Policy and Practice. Volume 55, September 2013, Pages 56–71

Yu, N., de Jong, M., Storm, S., & Mi, J. (2013). Spatial spillover effects of transport infrastructure: Evidence from Chinese regions. *Journal of Transport Geography, 28*(0), 56-66.

APPENDIX A: TABLES

Table 1: Descriptive Statistic for the Dependent and the Independent Varibles from year 1993 to year 2012

Variable:	Mean	Std. Dev.	Min	Max
(Urban-Rural)IncomeGap (%)	292.61	69.84	145.63	560.48
Railway density (km/km^2)	0.01685	0.0156	0	0.0796
Highway density (km/km^2)	0.4776	0.3961	0	1.9841
Transportation density(km.km^2)	0.4945	0.4060	0	2.0635
GDP per capita (1000RMB)	3.2703	2.8827	0.4504	16.0726
Average years in school	7.564	1.3572	2.1807	11.8363
Degree of openness (%)	30.15	39.41	0	220.295

Notes: The table includes percentages, means, and standard deviations, minimum and maximum for all variables. Data are from the 1993 to 2013 China Statistical Yearbooks. Numbers are calculated based on provincial data. There are 31 provinces included in our sample, Taiwan, Macau, and Hong Kong are not included due to the lack of data. Income gap is cumulated using the ratio of average disposable income in urban areas over net income for rural areas, and it is in percentage form. Degree of openness is measured using the ratio of total value of imports and exports of a province to the total GDP in that province.

Table 2: Results of Regression of Income Gap on Transportation density and Other Characteristics Using 1993 to 2012 panel data from China Statistic Year Books

Dependent variable: Urban-Rural Income Gap; 620 observations				
		OLS	0-1 weight	Inverse-distance weight
Regressor	(1)	(2)	(3)	(4)
transportation density	-15.38**	9.75	-14.788***	6.66
	(7.18)	(10.25)	(5.79)	(9.04)
Neighbor's density		-39.69***		
		(11.63)		
openness	24.758**	18.628**	-21.926***	-55.127***
	(7.642)	(7.78)	(4.370)	(6.134)
average years in school	-34.67***	-35.79***	-30.108***	-37.329***
	(2.17)	(2.17)	(1.661)	(2.022)
gdp	-27.33***	-25.53***	17.72***	18.078***
	(3.38)	(3.39)	(1.89)	(2.548)
gdpsq	1.496***	1.43***	-0.60***	-0.658
	(.184)	(0.18)	(0.13)	(0.165)
ρ			0.138***	14.357 ***
			(.00514)	(1.039)
λ			-0.115***	-12.392***
			(.00844)	(1.507)
Tests statistics and p-values				
LR Test :ρ=λ=0			982.0609	239.63
p-value			(0.000)	(0.000)
Wald Test			810.051	451.1285
p-value			(0.000)	(0.000)
F-Test	40.73	40.27	162.01	90.226
p-value	(0.000)	(0.000)	(0.000)	(0.000)
LM Error(Robust)			900.8763	3322.715
p-value			(0.000)	(0.000)
LM Lag (Robust)			49.7391	4.572
p-value			(0.000)	(0.984)
AIC			0.2141	0.286
Adj-R^2	0.151	0.147	0.543	0.389

Note: standard deviation in parenthesis
*** denotes statistical significance at the one percent, ** at the five percent, and * at the ten percent level

Table 3 Moran's I for Urban Rural Income Gap from 1993-2012

	1993-2007 Moran's I for Urban-Rural Income Gap					
	0-1 spatial weight matrix			Invers-distance spatial weight matrix		
Year	Moran's I	z	P-value	Moran's I	z	P-value
1993	0.236	9.655	0.000	0.264	7.267	0.000
1994	0.175	7.477	0.000	0.189	5.445	0.000
1995	0.215	23.661	0.000	0.277	7.604	0.000
1996	0.208	30.191	0.000	0.232	6.651	0.000
1997	0.199	35.797	0.000	0.213	6.346	0.000
1998	0.185	8.084	0.000	0.197	6.017	0.000
1999	0.191	8.268	0.000	0.210	6.289	0.000
2000	0.186	8.126	0.000	0.212	6.415	0.000
2001	0.199	8.536	0.000	0.230	6.755	0.000
2002	0.197	8.443	0.000	0.236	6.858	0.000
2003	0.187	8.007	0.000	0.230	6.598	0.000
2004	0.190	8.086	0.000	0.234	6.642	0.000
2005	0.187	7.948	0.000	0.234	6.595	0.000
2006	0.184	7.828	0.000	0.252	7.032	0.000
2007	0.198	8.314	0.000	0.257	7.150	0.000
2008	0.198	8.299	0.000	0.258	7.133	0.000
2009	0.189	7.980	0.000	0.252	7.014	0.000
2010	0.182	20.259	0.000	0.255	7.085	0.000
2011	0.184	7.807	0.000	0.242	6.777	0.000
2012	0.179	7.661	0.000	0.240	6.722	0.000

Table 4: Beta, Total, Direct and Indirect: Linear Marginal Effects of All dependent Variables on the Urban Rural Income Gap from year 1993 to year 2012

Variable	Beta(B)	Total	Direct	Indirect
Transportation density	-14.79	-13.02	-5.18	-7.84
GDP	17.72	15.60	6.20	9.40
GDP-sq	-0.60	-0.53	-0.21	-0.31
Average years in school	-30.11	-26.51	-10.54	-15.97
Openness	-21.9262	-19.3065	-7.6778	-11.6287

Note: The table reports the summary of marginal effect, total effect, direct effect and indirect effect of transportation density on the urban rural income gap. Regression controls for GDP per capita, square of GDP (measurement unit is 1000RMB), and average years in school in provincial level, and degree of openness.

Table 5: Bata, Total, Direct and Indirect: Linear Elasticity of All dependent Variables on the Urban Rural Income Gap from year 1993 to year 2012

Variable	Beta(ES)	Total	Direct	Indirect
Transportation density	-0.0250	-0.0220	-0.0088	-0.0133
GDP	0.198	0.1744	0.0693	0.105
GDP-sq	-0.0390	-0.0344	- 0.0137	- 0.0207
Average years in school	-0.7783	-0.6853	-0.2725	-0.4128
openness	-0.0226	-0.0199	-0.0079	-0.0120

Note: The table reports the summary of elasticity, total elasticity, direct elasticity and indirect elasticity of transportation density on the urban rural income gap. Regression controls for GDP per capita, square of GDP (measurement unit is 1000RMB), and average years in school in provincial level, and degree of openness.

Table 6: Results of Regreesion of Income Gap on Railway/Highway density and Other Characteristics Using 1993 to 2012 panel data from China Statistic Year Books

Dependent variable: Urban-Rural Income Gap; 620 observations				
	OLS		0-1 weight	Inverse-distance
Regressor	(1)	(2)	(3)	(4)
railway density	273.11	976.99***	-479.49***	-983.46***
	(201.69)	(257.64)	(170.00)	(206.8)
highway density	-19.20**	-13.72	-10.72*	10.40
	(7.66)	(11.57)	(5.995)	(8.89)
neighbors' railway		-1382.87***		
		(353.03)		
neighbors' highway		-4.70		
		(15.00)		
openness	24.385***	17.636**	-17.189***	-40.775***
	(7. 639)	(7.69)	(6.069)	(6.715)
average years in school	-36.21***	-36.94***	-27.41***	-31.4***
	(2.41)	(2.41)	(1.912)	(2.332)
gdp	-27.61***	-24.46***	15.65***	14.301***
	(3.38)	(3.38)	(2.02)	(2.618)
gdpsq	1.47***	1.31	-0.446***	-0.363***
	(0.185)	(0.18)	(0.142)	(0.174)
ρ			0 .135***	14.209***
			(.005)	(1.056)
λ			-0.1160286 ***	-12.711***
			(.0086447)	(1.646)
Tests statistics and p-values				
LR Test :ρ=λ=0			801.439	239.843
p-value			(0.000)	(0.000)
Wald Test			989.6395	690.985
p-value			(0.000)	(0.000)
F-Test(all coeffiencts =0)	39.25	38.9	164.9399	115.164
p-value	(0.000)	(0.000)	(0.000)	(0.000)
LM Error(Robust)			824.3642	2969.321
p-value			(0.000)	(0.000)
LM Lag (Robust)			36.073	6.795
p-value			(0.000)	(0.813)
AIC			0.1905	0.2342
Adj-R^2	0.141	0.157	0.594	0.501

Note: standard deviation in parenthesis

*** denotes statistical significance at the one percent, ** at the five percent, and * at the ten percent level

Table 7: Beta, Total, Direct and Indirect : Linear Marginal Effect of All Dependent Variables on the Urban Rural Income Gap from year 1993 to year 2012

Variable	Beta(B)	Total	Direct	Indirect
Railway density	-479.49	-425.89	-176.00	-249.89
Highway density	-10.72	-9.52	- 3.93	-5.59
GDP	15.64	13.90	5.74	8.15
GDP-sq	-0.446	-0.39	-0.1638	-0.2326
Average years in school	-27.41	-24.35	-10.06	-14.29
openness	-17.1886	-15.2673	-6.3093	-8.9580

Note: The table reports the summary of marginal effect, total effect, direct effect and indirect effect of railway and highway density on the urban rural income gap. Regression controls for GDP per capita, square of GDP, and average years in school in provincial level, and degree of openness.

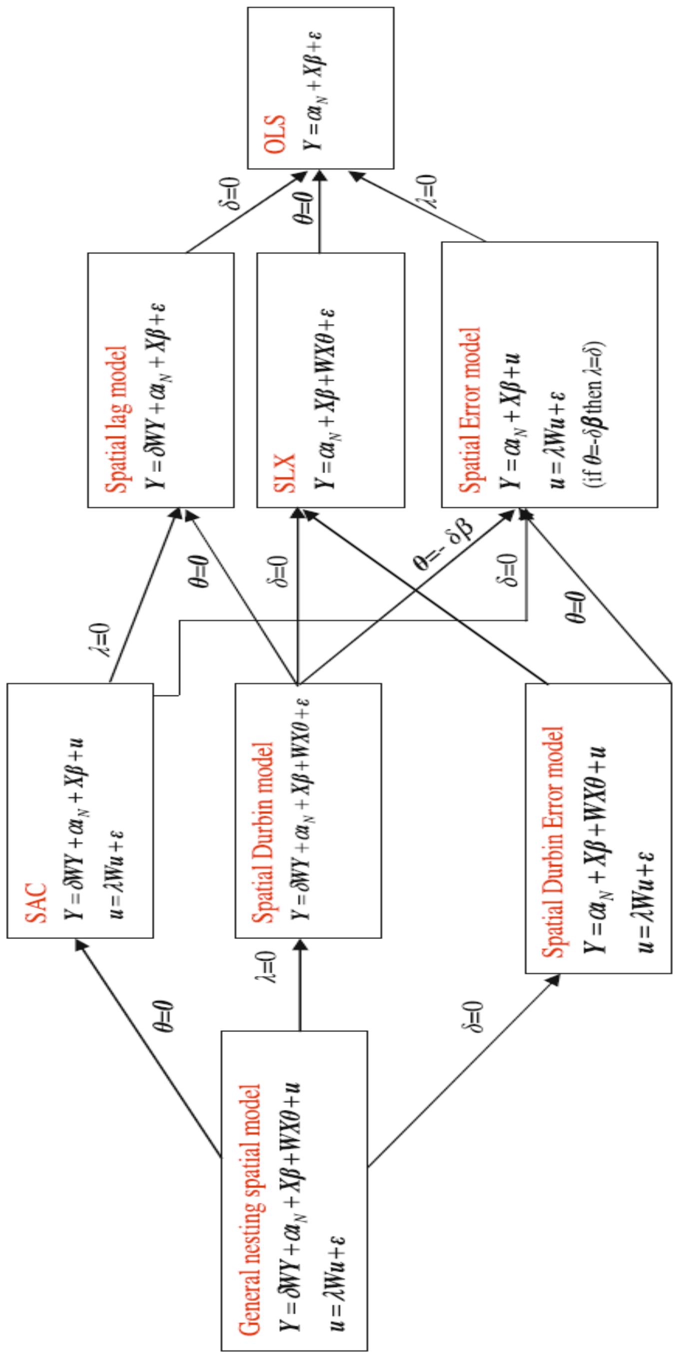
Table 8: Beta, Total, Direct and Indirect: Linear Elasticity of All Dependent Variables on the Urban Rural Income Gap from year 1993 to year 2012

Variable	Beta(Es)	Total	Direct	Indirect
Railway density	-0.0276	-0.0245	-0.0101	-0.0144
Highway density	-0.0175	-0.0155	-0.0064	-0.0091
GDP	0.1749	0.1554	0.0642	0.0912
GDP-sq	-0.0290	-0.0257	-0.0106	-0.0151
Average years in school	-0.7086	-0.6294	-0.2601	-0.3693
openness	-0.0177	-0.0157	-0.0065	-0.0092

Note: The table reports the summary of elasticity, total elasticity, direct elasticity and indirect elasticity of railway and highway density on the urban rural income gap. Regression controls for GDP per capita, square of GDP, and average years in school in provincial level, and degree of openness.

Appendix B: Figures

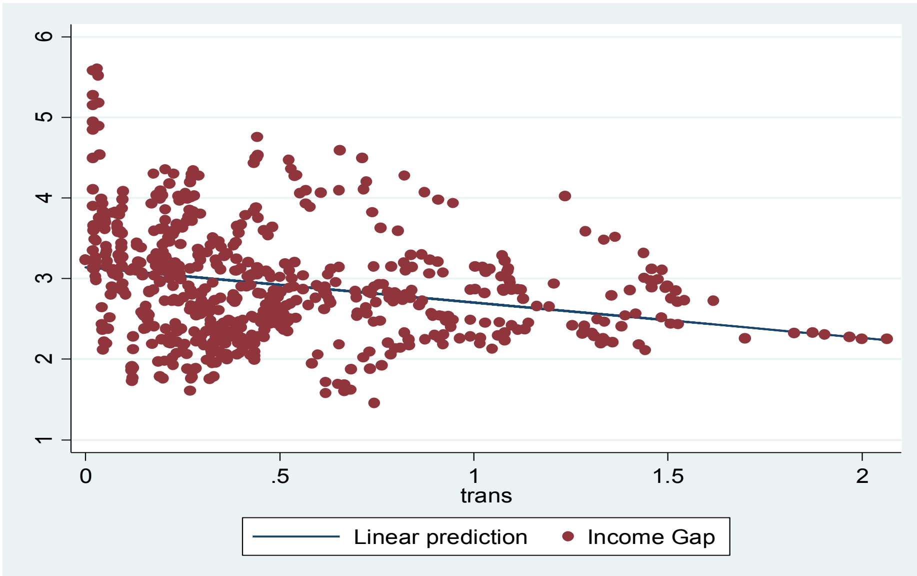
Figure 1: The Relationship between Different Spatial Dependence Models



Notes: In our case, we use "ρ" to represent "δ" in our spatial models.

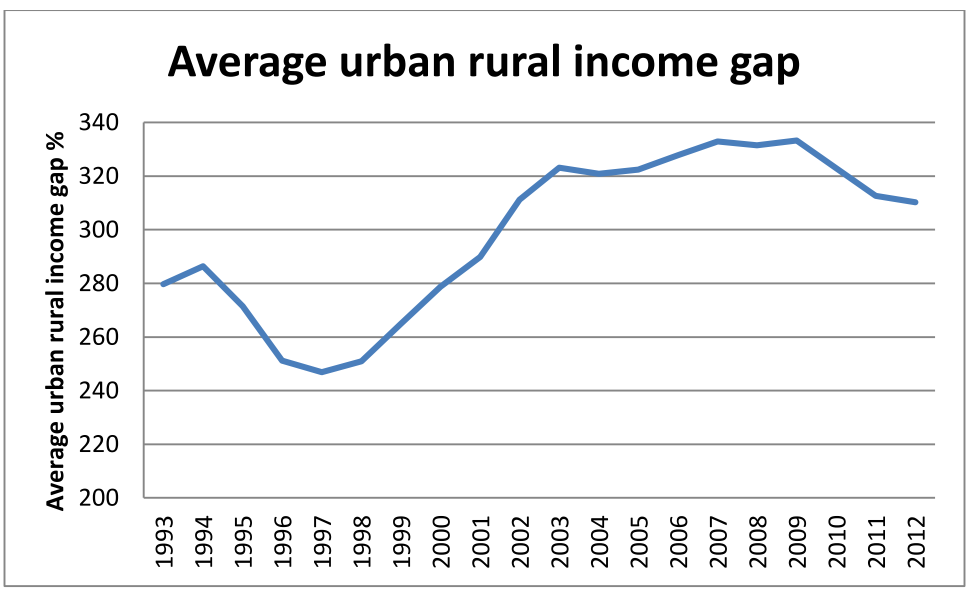
Source: Based on Halleck Vega and Elhorst (2012) as cited in Elhorst (2013): Spatial Econometrics: From cross-sectional data to spatial panels, page 9

Figure 2: Scatter Plot of Urban-Rural Income Gap Against Transportation Density



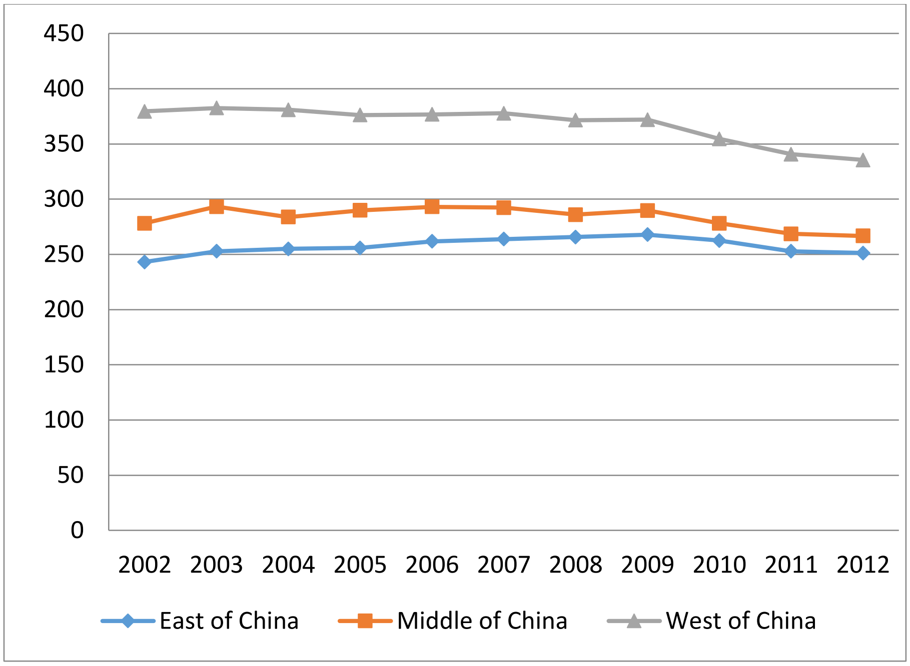
Source: 1993-2013 China Statistical Yearbook

Figure 3: Average Urban-Rural Income Gap in China from 1993 to 2012



Source: 1993-2013 China Statistical Yearbook

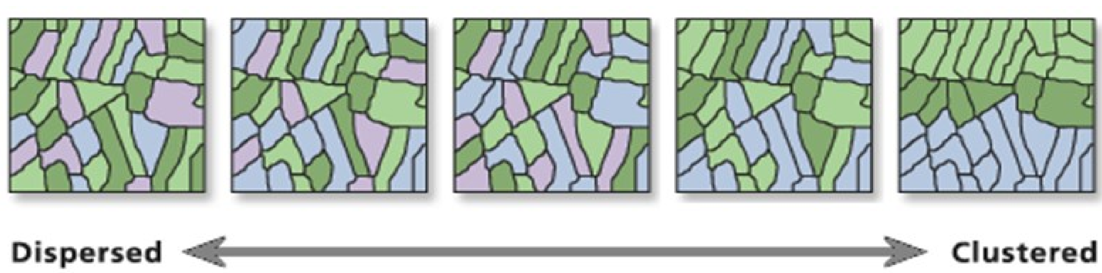
Figure 4: Urban-Rural Income Gap In China by Region from 2002 to 2012



Source: 1993-2013 China Statistical Yearbook

Note: The vertical axis represents the urban rural income gap measured in percentage.

Figure 5: Illustration of Spatial Autocorrelation Patterns

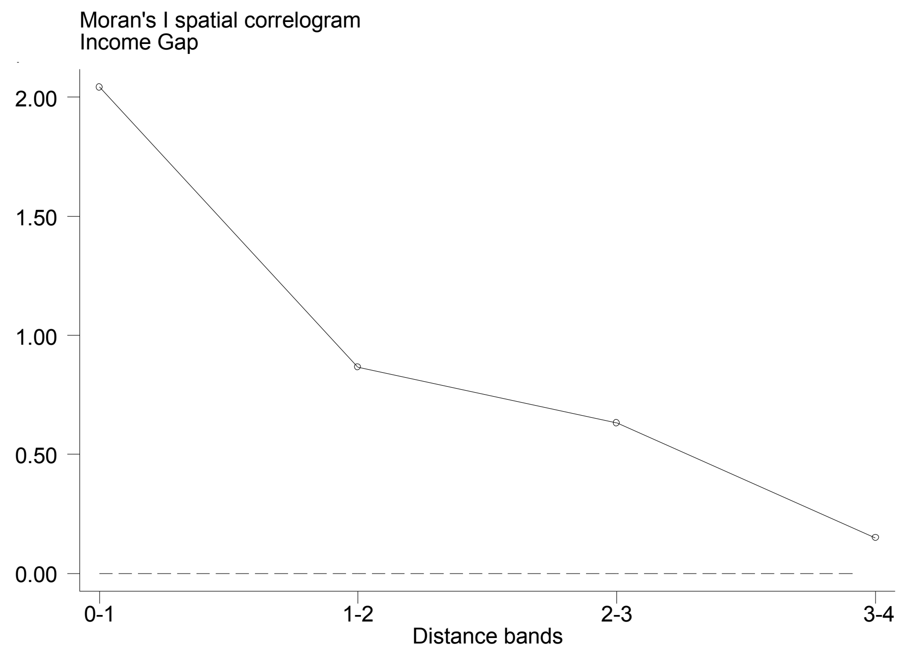


From online:

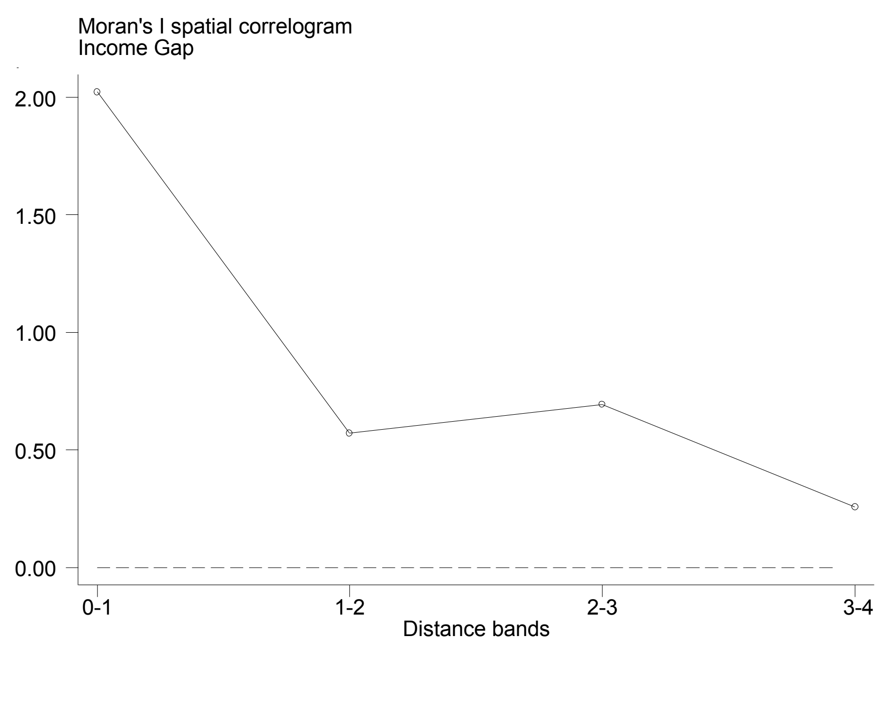
(http://edndoc.esri.com/arcobjects/9.2/NET/shared/geoprocessing/spatial_statistics_tools/spatial_autocorrelation_morans_i_spatial_statistics_.htm)

Figure 6: Moran's I Spatial Corelogram for year 1993 and year 2012

Year 1993



Year 2012



Source: Produced by STATA based on the data from China Statistical Yearbook 1994 and 2013