Subnational price differences in large emerging countries[1]

Yingchun Zhang *
Dongbei University of Finance and Economics, Dalian, China– zhangyingchun@dufe.edu.cn

1. Introduction

The International Comparison Program (ICP) is now organized by the World Bank (WB). ICP has become not only the largest international statistical program in the world but also the most complex.It aims to measure the real size of the world economy. The latest comparison is the 2011 ICP, which includes 199 economies. Each country must collect a large number of price quotations for each of the ICP product lists. Especially for some large developing/emerging countries. as we know, there is economic disequilibrium in different cities/locations/regions in these countries; thus, there are spatial price attributes. We do not know how to deal with the spatial price attributes when we obtain the national average, nor do we know how to identify them. This essay focuses on the identification of the spatial price attributes using apple prices in China.

The remainder of this essay includes discussions of the methodology and data (part 2),empirical process (part 3), interpretation of the results (part 4), conclusions(part5), and problems to be solved (part 6).

2. Methodology and Data

In many countries, especially large developing and emerging countries, economy disequilibrium is common within a country. Thus, the prices for the items on the ICP product lists do not have the same distributions in different pricing cities/locations. Every pricing city/locationdoes not have the same importance to the national average, and they thus should not be given same weight. In such cases, the arithmetic average is not perfect. Ideally, weight should be given according to the significance of the spatial price characteristics. We must study spatial heterogeneity and spatial agglomeration in order to determine the spatial price characteristics. These can be transformed to analyze spatial autocorrelation.

2.1 Spatial Autocorrelation

Spatial autocorrelation is a method of spatial data analysis used to study the potential interdependence of the observation data of some variable within certain region. That is, it is used to estimate whether the correlation between the observation at a position and the observation at an adjacent position are significant and to what degree. Spatial autocorrelation is used in many fields, including ecology, genetics, epidemiology, biology, and criminology.

2.2 Index of measuring spatial autocorrelation

Spatial autocorrelation includes global spatial autocorrelation and local spatial autocorrelation. There are many indexes that measure global and local spatial autocorrelation. The most commonly used include Moran's *I*, Geary's *C*, and the Getis–Ord *G*. This essay uses Moran's *I* because it has better statistical properties.

2.2.1 Global Moran's *I*

Let n denote the number of positions. x_i and x_j are the observations for positions i and j, respectively.w_{ij} is the adjacent degree between positions i and j. When i and j are adjacent, W_{ij}=1; otherwise, W_{ij}=0. Let I denote the index of Moran's I. The formula is as follows:

1 * This study was supported by National Social Science Major Projects: The Study of Theory and Practice of China's Full Participation in the ICP (Grant No.: 13&ZD171) and General Project of National Social Science Fund: The Study of International Comparison Methods on Non-tradable Products (Grant No.: 14BTJ001).

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n}\sum_{j\neq i}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{S^2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \qquad (1)$$

z is computed to measure spatial autocorrelation. If z is greater than zero and significant, there is positive spatial autocorrelation, which means that similar observations tend toward spatial agglomeration.

If z is less than zero and significant, there is negative spatial autocorrelation, which means that similar observations tend toward spatial heterogeneity. If z is zero, the observations are independent. The formula for z is given in (2).$E(I)$ is the expected Moran's I. VAR(I) is the variance of Moran's I.

$$z = (I - E(I))/\sqrt{VAR(I)} \qquad (2)$$

For the prices collected in every country in the ICP, if there is positive or negative spatial autocorrelation, the prices have certain spatial characteristics. These should be considered as subnational price differences and should be introduced into the calculation of the national annual average prices. If z is zero, the prices do not have obvious global spatial characteristics. If there are spatial characteristics, we would have to study the local spatial autocorrelation.

2.2.2 LISA

Local indicators of spatial association (LISA)are used to test the local spatial autocorrelation—specifically, to estimate the correlation between position i and the other positions. The formula is given in (3).The judgment rule for spatial autocorrelation makes it similar to Moran's I.

$$I_i = \frac{n^2}{\sum_i\sum_j w_{ij}} \times \frac{(x_i - \bar{x})\sum_j w_{ij}(x_j - \bar{x})}{\sum_j (x_j - \bar{x})^2} \qquad (3)$$

2.3 Data

Every economy participating in the ICP needs to collect a large number of prices for goods and services. These are taken from the final consumption expenditure and gross fixed capital formation. There are many items on the product lists—some countries have more than 1,000—and this essay could not and need not examine all of them. We selected apples, a common product. Its prices are relatively easy to collect from nearly every city/location in an entire country, which makes it suitable for spatial econometric analysis. The price data come from 31 provinces in China (Table 1).[2]

Table 1.Apple prices in 31 provinces in China (￥)

District	Price	District	Price	District	Price
Beijing	6.76	Zhejiang	17.5	Hainan	5.29
Tianjin	8.2	Anhui	6.5	Chongqing	12
Hebei	4.14	Fujian	5.1	Sichuan	13
Shanxi	9.88	Jiangxi	6.5	Guizhou	3.6
Neimenggu	5.07	Shandong	5.79	Yunnan	5.5
Liaoning	4.95	Henan	4.83	Xizang	3
Jilin	4.46	Hubei	8.2	Shanxi	8
Heilongjiang	4.87	Hunan	11.5	Gansu	12
Shanghai	11.3	Guangdong	10.26	Qinghai	4

2Chinese administrative divisions include provinces, autonomous regions, and municipalitiesdirectly under the central government. This essay callsthem "provinces" for short.

Jiangsu	10.97	Guangxi	9.63	Ningxia	3.7
-	-	-	-	Xinjiang	6.5

Data Source: http://www.21food.cn/

3. Empirical Process

3.1 Moran's *I*

The software used here is GeoDa. The geospatial base map is China. The Moran's *I* for the global spatial autocorrelation is shown in Figure1. The value of the index is 0.2964.This is greater than zero; thus, the apple prices in the 31 Chinese provinces have a positive spatial correlation.

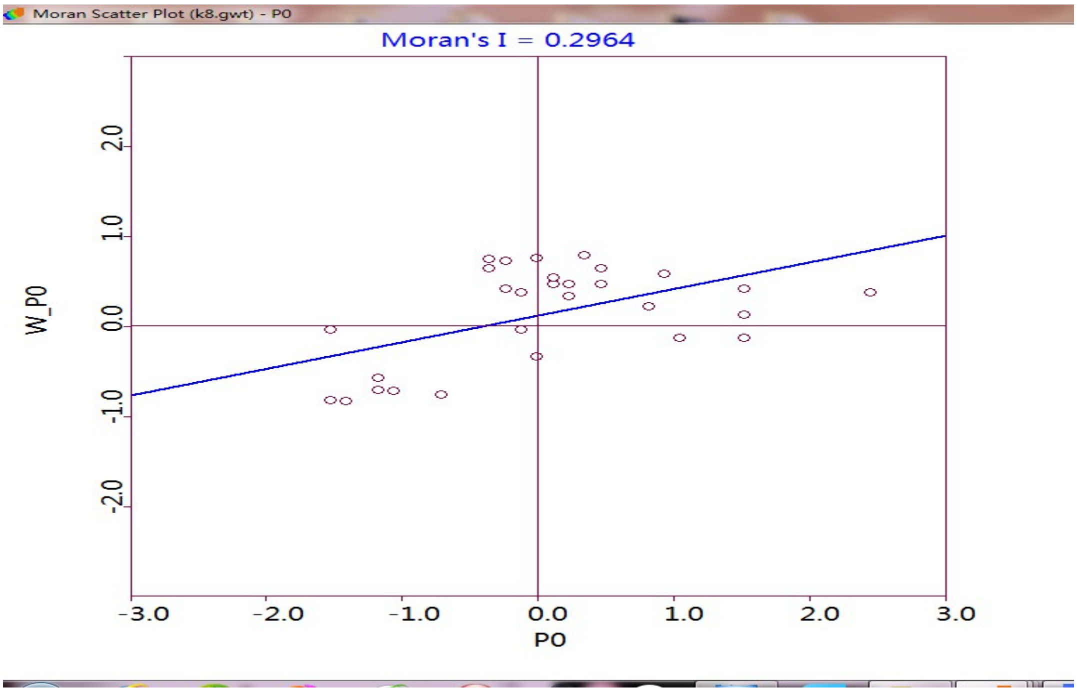

Figure1. Moran's *I*

3.2 LISA map

The LISA map for the local spatial autocorrelation is shown in Figure2. There are obvious spatial price differences. The 31 provinces can be divided into four groups. The first group consists of provinces with high prices surrounded by other provinces with high prices: Sichuan, Chongqing, Hubei, Hunan, Guangxi, Guangdong, and Zhejiang. The second group consists of provinces with low prices surrounded by other provinces with low prices: Neimenggu, Heilongjiang, Jilin, Liaoning, Hebei, and Beijing. The third group consists of provinces with low prices surrounded by other provinces with high prices: Anhui, Jiangxi, and Fujian. The fourth group consists of the rest 15 provinces, which have no clear rules.

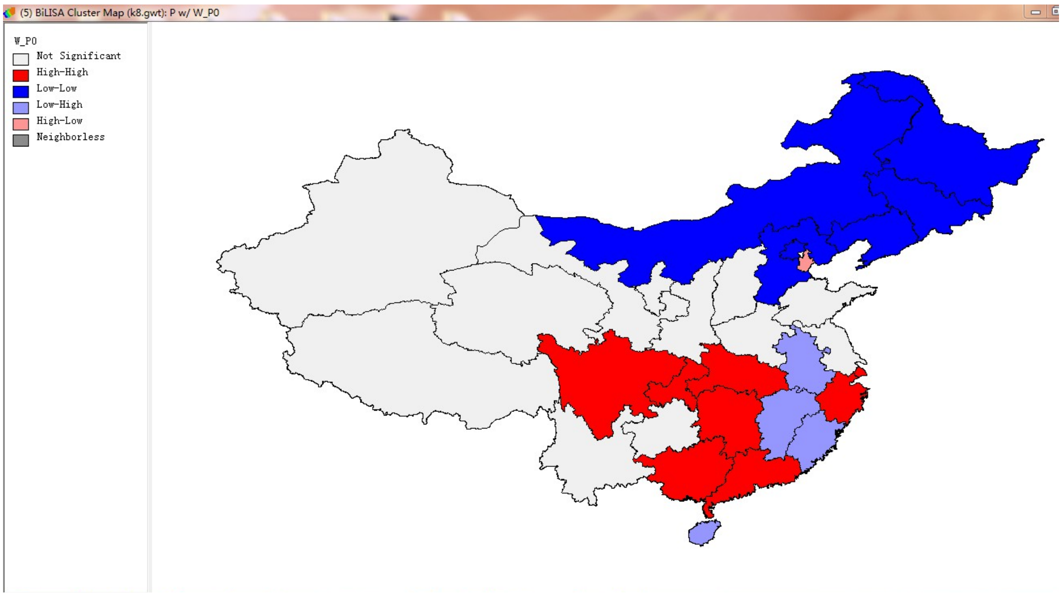

Figure2. The LISA map

4.Interpretation of the results

4.1 Interpretation of Moran's *I*

As we know, the value of Moran's *I* is between -1 and 1. If Moran's *I* is greater than 0, there is positive spatial correlation. If Moran's *I* is nearly 1, the observations in the spatial cluster have similar characteristics; that is, high observations are adjacent to other high observations and low observations are adjacent to other low observations. If Moran's *I* is less than 0, there is negative spatial correlation. If Moran's *I* is nearly -1, the observations in the spatial cluster have dissimilar characteristics; that is, high observations are adjacent to low observations, and vice versa. If Moran's *I* is nearly 0, there is no spatial correlation; that is, the observations are spatially independent.

The value of Moran's *I* in Figure1 is 0.2946. Obviously, it is not 0; therefore, the apple prices in the 31 provinces are not spatially independent: there is positive spatial correlation. This suggests that the simple arithmetic average is not the ideal value. We should look for another average that could reflect the spatial characteristics. Perhaps a weighted average would be more ideal. The core question is how to assign the weight. These issues we left to our future work.

4.2 Interpretation of LISA

Moran's *I* indicates that the prices are not spatially independent—there is positive spatial correlation. The next step is to determine the spatial heterogeneity or spatial homogeneity. LISA is used to solve this problem. Figure2 shows the four groups of provinces. The first and second groups have spatial homogeneity, while the third group has spatial heterogeneity. The fourth group is spatially independent. These spatial characteristics should not be neglected at all, as they can affect the accuracy of the national average.

The first group is comprised of seven provinces with high prices surrounded by others with high prices. These seven provinces have the highest prices in the entire country. They contribute substantially to increasing the national average price, as the average price for this group would be higher than the national average price. The second group is comprised of six provinces with low prices surrounded by others with low prices. These six provinces have the lowest prices in the entire country. They contribute substantially to decreasing the national average price, as the average price for this group would be less than the national average price. The third group is comprised of three provinces with low prices that are surrounded by others with high prices. They have spatial heterogeneity. The fourth group has no obvious spatial characteristics; these provinces appear to be spatially independent.

Table 2 shows the average prices of the four groups and the whole country. The data support the above findings. The average for the first group is far larger than that for the whole country, while the second group's average is much lower. The fourth average is 6.84; these provinces are spatially independent, and the fourth average is closest to that for the whole country. Could we have a bold guess that the average for the provinces with spatially independent can replace that for the whole country. This seems wrong because many provinces are not included; nevertheless, it is correct. However, this could be interpreted as being so because the extreme values have been eliminated. Whether this is practicable is left to future research.

Table2. Average prices based on the spatial characteristics (￥)

Region	First Group	Second Group	Third Group	Fourth Group	Whole Country
Number of provinces	7	6	3	15	31
Average price	11.72	5.04	6.03	6.84	7.52

5. Conclusions

5.1 The prices have obvious spatial characteristics

The apple prices in the 31 Chinese provinces have obvious spatial characteristics. The prices are not spatially independent; they have a positive spatial correlation. Further, there are 13 provinces with substantially higher or lower prices, comprising more than one third of the whole country. To obtain a more accurate national average, the 13 provinces may require further attention. This is left to future research.

5.2 ICP's national average prices have spatial attributes

According to the WB handbook, the ICP's national average prices are for the different cities/locations in a country. If a country consists of many cities, prices should be collected from every city. In some cases, not all of the cities could satisfy the demand for data. This would not have a significant impact unless the city was a pillar of the country, or the opposite, the city is dispensable. From the perspective of spatial attributes, the data in these cities are substantially higher or lower. This is common in some large emerging countries (for example, the first and second groups discussed above).

5.3 ICP's arithmetic average needs to improve

ICP's national annual average prices are calculated using the arithmetic average method. The average prices should ideally cover the whole country, representing both rural and urban regions as well as different geographic regions for larger countries. However, different regions with different prices have different effects on the average. Many cities/locations/regions have different speeds and qualities of economic development, especially in large emerging countries. We do not think the arithmetic average could effectively cover all such regions. We should calculate other averages in addition to the spatial price characteristics given in this essay to improve the ICP's average.

6.Problems to be solved

1. Could the national average be replaced by the average of the cities/locations/regions with spatially independent attributes?
2. Are there any other methods that can be used to study the spatial characteristics of the prices?
3. Could spatial sampling be used to reflect the spatial characteristics?
4. How can the national average be weighted according to the spatial characteristics of the prices?
5. Do all the products in the ICP's list have spatial characteristics, and if so, how can they be identified?

Conference

1. World Bank. 2013. *Measuring the Real Size of the World Economy: The Framework , Methodology, and Results of the International Comparison Program- ICP*. Washington, DC: World Bank. DOL:10.1596/978-0-8213-9728-2.

2. European Union/OECD. 2012. *Eurostat-OECD Methodological Manual on Purchasing Power Parities*. Publications office of the European Union. DOL: 10.2785/33942
3. Ping Jiang etc. 2014.3. System of National Accounts and Its Implementation in China. China Renmin University Press.
4. Yingchun Zhang. 2013. New Development in Statistical Studies of the Global Economy and Its Enlightments for China. Social Sciences Academic Press(China).