Regional Specialization: A Measure Method and the Trends in China

Lu, Zheng and Flegg, A.Tony and Deng, Xiang

School of Economics, Sichuan University, P.R.China, Department of Economics, Bristol Business School, University of the West of England, UK, Department of Economics, Sabanci University, Turkey

1 October 2011

Regional Specialization: A Measure Method and the Trends in China

Zheng Lu[a*], Tony Flegg[b] and Xiang Deng[c]

a. *School of Economics, Sichuan University, Chengdu, P.R.China; Department of Economics, Sabanci University, Istanbul, Turkey.*

b. *Department of Economics, Bristol Business School, University of the West of England, Bristol, UK.*

c. *School of Economics, Sichuan University, Chengdu, P.R.China.*

Abstract

This paper elaborates on a method of measuring regional specialization and examines the trend of regional specialization in China, 1987 - 2007. It constructs a simple coefficient incorporating the effect of regional industrial scale, based on location quotients, and then measures the regional specialization of China using official statistical data. The results indicate a remarkable increase in China's overall regional specialization during this time, as well as obvious regional and industrial differences, i.e., that the regional specialization of eastern coastal China is relatively less than that of the inland. Findings further demonstrate that special-resource-dependent industries are concentrated in regions with resource endowment, whereas industries with strong technical barriers are mainly located in regions with strong research and innovation ability.

Key Words: Regional Specialization; Location Quotients; China

JEL Classification: C69; P25; R11; R12

1. Introduction

Geographic distribution of economic activities is an important research field of economics. From 1990s, research on industrial concentration and regional specialization developed very fast coming with the development of new economic

---

We are grateful to the research funding of National Natural Science Foundation of China(Project Number: 70773078).

[*]Corresponding author: Zheng Lu, School of Economics, Sichuan University, Wangjiang Rd. 29, Chengdu, 610064, P. R. China, E-mail: zluecon@gmail.com.

geography theory, agglomeration economics and spatial economics. And meanwhile, empirical studies also made some progress due to the improving of data availability and quality, studies on U.S. by Kim (1995), Ellison and Glaeser (1997), studies on North America by Holmes and Stevens (2004), and studies on EU by Amiti (1998) and Brülhart (2001) are the representative contributions.

Another important reason for the sustained development of empirical studies is the development of measure method for industrial geographic distribution and regional specialization. Methods such as Hoover coefficient, Gini coefficient, Krugman's specialization coefficient and Ellison-Glaeser coefficients are widely used for measuring the level and tendency of regional specialization. These methods provided necessary technological means for empirical studies.

China's economic reforms and rapid growth over the past few decades has attracted a great deal of attention from economists, and many economists have begun to study the problem of industrial geographic distribution and regional specialization in China. Studies of Young (2000) and Poncet (2003) revealed that China's regional specialization declined in 1980s and 1990s because of China's descending domestic integration in these periods, but others studies based different measure methods argued that regional specialization in China increased after the late 1980s (e.g. Brun and Renard, 2002; Naughton, 2003; Bai et al., 2004; Liang and Xu, 2005).

Our focus in this paper is on the measure method of regional specialization and China's regional specialization in the past two decades. We attempt to construct a new coefficient after analyzing current measuring methods and then use it to measure regional specialization in China. The rest of this paper is organized as follows: Part 2 reviews coefficients of measuring regional specialization; Part 3 discusses the methodology of measuring regional specialization; Part 4 analyzes the historical trend of regional specialization in China from 1987 to 2007; and the last part is conclusion.

2. Review of Literature

There are two paths to reflect the general level of regional specialization, one is to directly measure specialization level of a region, another which can reflect general regional specialization indirectly is to measure localization level of a industry, thus a coefficient of measure regional specialization always has two expressions, one for regions and another for industries. Coefficients such as location quotient (LQ), the Hoover Coefficient (Hoover, 1936), Locational Gini Coefficient (Krugman, 1991), and Ellison-Glaeser Coefficients (Ellison and Glaeser, 1994, 1997) are widely used to analyze industrial geographic distribution and regional specialization. Some Chinese researchers also proposed improved coefficients, for example, Lu and Tao (2005) constructed β_r coefficient based on Hoover and Eellison-Claeser's coefficient, β_r coefficient considers the differences of regional economic size.

2.1. Coefficients of Measuring Specialization

At the first, we provide some indicator assumptions in order to make following discussions much more easier and clearer. Assume that a geographic whole which we will regard it as a country is divided into M regions, and total number of industry is N. E_{ri}, NE_i, TRE and TNE are, respectively, employment of industry i in region r, employment of industry i in the entire country($NE_i = \sum_{r=1}^{M} E_{ri}$), total employment of all industries in region r($TRE = \sum_{i=1}^{N} E_{ri}$), and total employment of all industries in entire country($TNE = \sum_{i=1}^{N} \sum_{r=1}^{M} E_{ri}$).

Hoover localization coefficient (Hoover, 1936) was used to measure industries localization in any region. Firstly, for any region we calculate the employment share of industry i in region r in all industries employment of region r(E_{ri}/TRE) and the employment share of industry i in all industries of the entire country(NE_i/TNE). Then sort these two ratios in descending order and calculate their cumulative value. And finally, draw these two cumulative ratios on the vertical axis and abscissa axis separately, thus we obtain a localization curve of industry i. Hoover localization coefficient was defined as the ratio of area between localization curve and 45-degree line divided by area of the triangle. If interregional distribution of an industry is totally uniform, the coefficient equals to 0, and if an industry concentrated entirely in one region, the coefficient is 1. Krugman(1991) used a similar index which is called "Locational Gini Coefficients" to measure industry localization, the distinction is Krugman's index just equals to the area between localization curve and 45-degree line, thus its value interval is [0, 0.5].

Ellison and Glaeser (1994, 1997) used a easier index which is also known as Spatial Gini coefficient when they study the geographic concentration in U.S.. Assume that s_{ri} is employment share of industry i in region r in total employment of industry i, $s_{ri} = E_{ri}/NE_i$, s_r is employment share of region r in aggregate employment, $s_r = TRE/TNE$. Then Spatial Gini coefficient for industrial localization is defined by

$$G_i = \sum_r (s_{ri} - s_r)^2 \qquad (1)$$

If $G_i = 0$, it implies the industry distributes uniformly across regions, a greater value indicates a higher spatial concentration. Since it emphasizes departures of a given industry's employment, spatial Gini coefficient is of economic interest, and moreover, it is easier to work with than Gini coefficient (Ellison and Glaeser, 1994). In order to measure specialization for a given region, we need to carry out a simple transformation and then calculate summation by industry. Let S_{ri} represents employment share of industry i in region r in all industries of region r, $S_{ri} = E_{ri}/TRE$, and $S_i = NE_i/TNE$ represents employment share of industry i in aggregate employment. Then spatial Gini coefficient for regional specialization can be defined by

$$G_r = \sum_{i=1}^{N}(S_{ri} - S_i)^2 \quad (2)$$

Another similar alternative method is industrial specialization index proposed by Krugman (1991). Industrial specialization index measure specialization through measuring industrial structure difference of inter-region, difference of industrial structure between region r_1 and r_2 can be measured by equation (3).

$$S_{r_1 r_2} = \sum_{i=1}^{N}\left|S_{r_1 i} - S_{r_2 i}\right| \quad (3)$$

If industry structure of region r_1 is the same as r_2, the index is equal to 0. If industry structure of region r_1 is completely irrelevant with r_2, index is going to be 2. Therefore, Krugman's specialization coefficient can reflect industry structure difference and regional specialization level.

As similar as Gini coefficients, Herfindahl index also are two kinds of expressions, one for measuring industrial geographic concentration (localization) and another is for regional specialization. Herfindahl index for industrial localization is defined by

$$H_i = \sum_r (s_{ri})^2 \quad (4)$$

For a given industry i, if every region get the same share, Herfindahl equals to $1/M$, and if the given industry only concentrates in one region, Herfindahl equals to 1. Following a similar transformation, Herfindahl index for regional specialization is defined by

$$H_r = \sum_{i=1}^{N}(S_{ri})^2 \quad (5)$$

Value interval of Herfindahl index for regional specialization is $[1/N,\ 1]$. If all industries in region r have equal shares, index is equal to $1/N$, and if region r specialized in only one industry, index equals 1.

However, Hoover coefficient, Locational Gini coefficient, Spatial Gini coefficient and Herfindahl index mentioned above have a common defect of missing the impact of firm location behavior, and as everyone knows, plant's distribution is an very important factor of industrial localization as well as regional specialization. Fortunately, Herfindahl method is always used to measure industrial concentration using firm's market share, thus if enterprise-level data is available, it enables us to construct a specialization index including plant's geographic distribution. Assume that there are m firms in industry i and market share of firm k is z_k, then Herfindahl index $H_i^* = \sum_k (z_k)^2$, $k = 1, 2, \cdots$. Ellison and Glaeser (1994,1997) proposed an index known as γ_i index for measuring industry agglomeration effect both including effect of firm size and region size, γ_i index mainly consists of G_i and H_i^*,

$$\gamma_i \equiv \frac{G_i - \left(1 - \sum_{r=1}^{M} s_r^2\right) H_i^*}{\left(1 - \sum_{r=1}^{M} s_r^2\right)\left(1 - H_i^*\right)} \tag{6}$$

Theories of Ellison and Glaeser (1994, 1997) provided some inspiration for solving similar problem in measuring the specialization of a region. In terms of the constructing method of γ_i index , Lu and Tao (2005) proposed β_r index for measuring regional specialization,

$$\beta_r \equiv \frac{G_r - \left(1 - \sum_{i=1}^{N} S_i^2\right) H_r^*}{\left(1 - \sum_{i=1}^{N} S_i^2\right)\left(1 - H_r^*\right)} \tag{7}$$

where $H_r^* = \sum_{k=1}^{K} (E_{rk} / \sum_{k=1}^{K} E_{rk})^2$ is Herfindahl index of region r, and E_{rk} is employment or output of firm k in region r, K is total number of firms in region r.

Based on these methods, there are many empirical studies for measuring China's regional specialization. Although some studies found that China's regional specialization declined in 1980s and 1990s because of China's descending domestic integration in these periods (Young, 2000; Poncet, 2003), most of studies argued that regional specialization in China increased after the late 1980s. Brun and Renard (2002) and Naughton (2003) proved an accentuation of regional specialization between the late 1980s and early 1990s. Bai et al. (2004) found that regional specialization in china from 1985 to 1997 reversed at 1988 using Hoover coefficient, it registered a significant increase after 1988. Using Krugman's specialization index, Liang and Xu (2005) found except for Hainan and Hubei, all other provinces experienced an increase in specialization from 1988 to 2001. Studies of Lu and Tao (2005) which made use of β_r and firm's data from 1998 to 2003, and studies of Fan (2007) which used the index constructed by himself from 1985 to 2004 also obtained same conclusions. Anyway, most of studies of China's regional specialization obtained similar conclusion: regional specialization has clearly increased since China began to implement its reformation and opening strategy .

2.2. LQ and FLQ

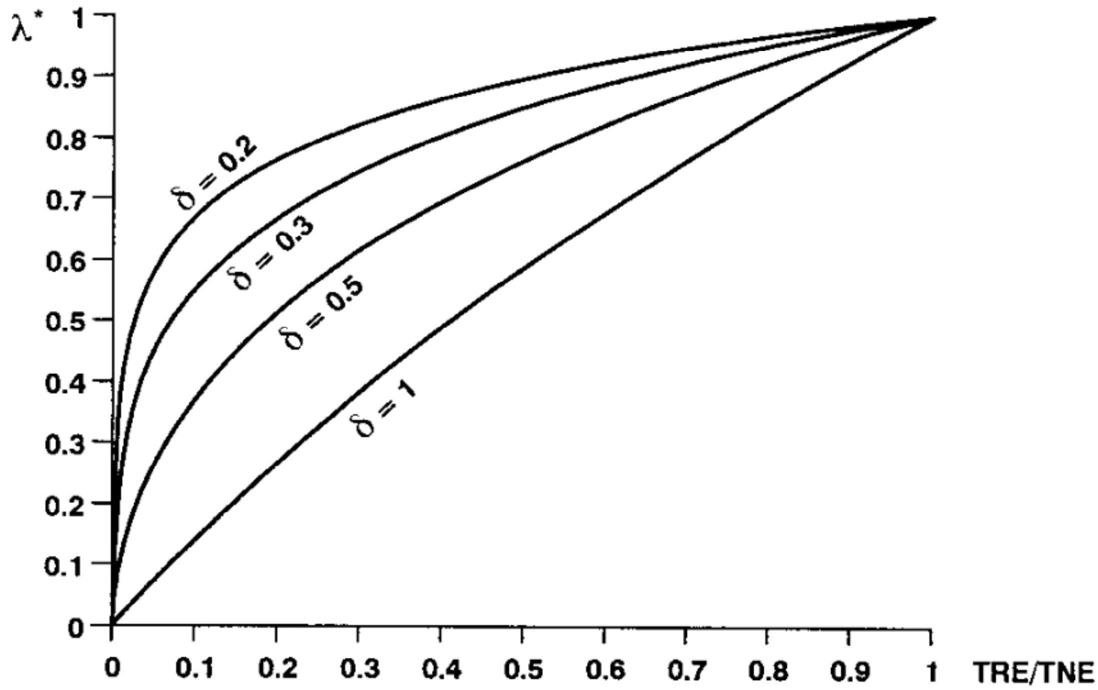
Compared with those methods mentioned above, LQ coefficient is a simple method for measuring industrial geographic concentration and, further, for specialization (Actually, all of aforementioned coefficients are related to LQ). LQ was presented by Haggett (1965) and used to analyze regional problems and also employed to measure the industrial geographic concentration and regional specialization. As defined by Flegg and Webber (1997), the simple LQ of industry i in region r of a country is,

$$SLQ_{ri} = \frac{E_{ri} / TRE_r}{NE_i / TNE} = \frac{E_{ri}}{NE_i} \times \frac{TNE}{TRE_r} \quad (8)$$

where RE_{ri}, NE_i, TRE and TNE still represent the same meaning defined previously. An important defect of SLQ is that it does not consider the effect of regional size which, as we knew, is an essential part of a coefficient for analyzing the interregional trade relation(Round, 1978). In order to overcome this problem, Flegg and Webber (1997) put forward a formula which involves the effect of region size,

$$FLQ_{ri} = SLQ_{ri} \times \lambda^*, \text{ where } \lambda^* = [log_2(1 + TRE_r/TNE)]^\delta \quad (9)$$

δ is sensitivity, δ is smaller, regional scalar λ^* becomes more concave, and FLQ and SLQ coincide while δ=0. Finally, FLQ involves not only the relative size of supplying and purchasing sector but also the regional size. As the meaning as LQ, FLQ also reflect the degree of industrial geographic concentration, it differs importantly in that FLQ considers the effect of regional size.



Source: Flegg & Webber(1997).

Figure 1: The Concave Property of λ^*

Comparing location quotients with Gini, Hoover, Herfindahl and β_r coefficients, we can find that, actually, all coefficients of measuring regional specialization are improved index based on LQ, their main component is employment shares. However, location quotients also reflect obvious defects: firstly, it cannot directly describe the trend of industrial geographic concentration or regional specialization in time dimension, they can only measure the concentration of an industry in a given region, in other words, the result is a three dimension dataset composed of time, region and industry; Secondly, both LQ and FLQ ignored the impact of industrial regional scale. Although coefficients of measuring regional specialization mentioned above overcome the first problem, they still inherit the second defect of location quotients. Moreover, it is also difficult to use these methods owing to the complicated computational process or the requirement of complete and high quality data. This paper thus attempts to find another simpler method to measure regional specialization based on LQ and FLQ.

3. Methodology

3.1. Location Quotient Considered Industrial Regional Scale

Since economic activity can be measured by not only employment but also output and earnings, now Let E represent general economic activity and it can be expressed by indicators such as employment, gross output, value added and so on, let r and i still represent region and industry, and M, N are total number of regions and industries, δ is the Flegg's sensitivity factor ($0 \leq \delta < 1$), further we define that S_{ri} is the E share of industry i in region r in total E of region r, S_i is the E share of industry i in total E, s_{ri} is the E share of industry i in region r in total E of industry i, and s_r is the E share of region r in total E. Then, the general expressions of LQ and FLQ are,

$$LQ_{ri} = \frac{E_{ri} \Big/ \left(\sum_{i=1}^{N} E_{ri}\right)}{\left(\sum_{r=1}^{M} E_{ri}\right) \Big/ \left(\sum_{i=1}^{N} \sum_{r=1}^{M} E_{ri}\right)} = \frac{S_{ri}}{S_i} \tag{10}$$

$$FLQ_{ri} = LQ_{ri} \times \left[log_2 \left(1 + \frac{\sum_{i=1}^{N} E_{ri}}{\sum_{i=1}^{N} \sum_{r=1}^{M} E_{ri}} \right) \right]^{\delta} = \frac{S_{ri}}{S_i} \times \left[log_2 \left(1 + s_r \right) \right]^{\delta} \tag{11}$$

Obviously, if we use FLQ to measure the industrial geographic concentration, the result differs from that obtaining when using LQ. For an industry in a region, FLQ may be less than LQ because of a smaller regional size (s_r), which means the specialization of the industry in this region measured by FLQ is less than that measured by LQ. For example, we computed the LQ and FLQ of the Beverage Manufacture in every one of the 31 provincial administrative regions in China in 2007 (considering the completeness of statistical data, we chose Gross Industrial Output Value as the indicator of measuring economic activity). Tibet obtained the highest LQ value 9.80, the second one is Sichuan province whose value is about 4.50, but the FLQ (δ=0.3) of this industry in Sichuan province(1.70) is much greater than that of Tibet(0.68). It's easy to find that the reason for this inversion is the effect of regional size, the Gross Industrial Output Value of all industries in Tibet is much less than that of Sichuan province. Actually, in 2007, the gross output value of the manufacture of beverages in Tibet is only 0.497 billion(0.1% of this industry in China), and it is 62.424 billion in Sichuan province (12.2% of this industry in China, ranking 1st), which indicates that in the macro perspective, the manufacture of beverages mainly concentrates in Sichuan province. Thus, we can conclude that as a method of measuring the industry concentration, FLQ is more objective and accurate than LQ.

But on the other hand, FLQ also reveals that it could underestimate or overestimate industrial geographic concentration. This problem can also be sufficiently explained by the comparison between Tibet and Sichuan province. The gross industrial output value in Tibet accounts for 0.01% of the total output value of

China, and the share of Sichuan province is 2.73%, then regional scalar λ^* are 0.071 and 0.377 respectively (δ=0.3), this means the adjustment to Tibet is much stronger than that to Sichuan province. If the industry of Tibet had a bigger share, this stronger adjustment by regional size would lead to underestimation of the concentration of this industry. On the contrary, for a region with a large size, adjustment by regional size may lead to overestimation of the concentration of those industries which have a very small share in entire country.

Generally, for any two regions r_1, r_2 and a given sensitivity factor, if $s_{r_1} < s_{r_2}$ and $s_{r_{1i}} \geq s_{r_{2i}}$, FLQ would underestimate the concentration of industry i in region r_1 in that regional scale of industry i in region r_1 is exact larger than that in region r_2. In the contrary, if $s_{r_1} \geq s_{r_2}$ and $s_{r_1 i} < s_{r_2 i}$, results of FLQ wound be reverse. However, in other cases, results depend on the differences of s_r and s_{ri} between two regions, for example, if s_{r_1} is only greater than s_{r_2} a little bit, while $s_{r_{1i}}$ is much more greater than $s_{r_{2i}}$, FLQ would underestimate the concentration of industry i in region r_1. Consequently, regional size(s_r) may not the best choice of constructing regional scalar, it involves lots of economy information, but at the same time, it internalized the effect of interregional differences in industry structure.

Table 1: Measure Results of FLQ in Different Cases

	$s_{r_1} \geq s_{r_2}$	$s_{r_1} < s_{r_2}$
$s_{r_1 i} \geq s_{r_2 i}$	Depends on differences of s_r and s_{ri} differences between two regions	Underestimates the concentration of industry i in r_1, and overestimates the concentration of industry i in r_2
$s_{r_1 i} < s_{r_2 i}$	Underestimates the concentration of industry i in r_2, and overestimates the concentration of industry i in r_1	Depends on differences of s_r and s_{ri} between two regions

Considering above problems, we can find that although relative regional size(s_r) is a good adjustment factor of location quotient, it is still too rough to measure industrial geographic distribution, and actually, in the review of Gini coefficient for concentration (G_i) and Hoover coefficient for concentration (H_i), we also can realize the importance of regional industrial scale (s_{ri}) when economists construct a coefficient for measuring industrial geographic concentration. Thus, we argued that the most important factor should be not regional size (s_r) but regional industrial scale (s_{ri}). So, in this paper, we use the industrial regional scale as an adjustment factor, then, the FLQ can be rewritten as following equation,

$$FLQ_{ri}^{\star} = LQ_{ri} \times \left[\log_2(1 + \frac{E_{ri}}{\sum_{r=1}^{M} E_{ri}}) \right]^{\delta} = \frac{S_{ri}}{S_i} \times \left[\log_2(1 + s_{ri}) \right]^{\delta} \quad (12)$$

3.2. Application of LQs in Measuring Regional Specialization

LQ is widely used to measure the level of industrial concentration, FLQ and $FLQ^{\star}$ are improved by regional size or industrial regional scale based on traditional location quotient, thus FLQ and $FLQ^{\star}$ also can reflect industrial geographic concentration. But the problem is, their values indicate the concentration of every industry in a region or the concentration of an industry in every region in one year, which can be expressed by a two dimension table or matrix, so it is problematic to analyze the general specialization of a country by using these methods. Firstly, LQ, FLQ and $FLQ^{\star}$ cannot analyze the geographic concentration of a certain industry at the national level. Secondly, they cannot analyze the overall specialization of a certain region, because there is not only one industry in a region and industries differ from each other in the concentration level. Thirdly, it is difficult to clearly express if the region and industry are divided into too small segments or in too many parts (Lu and Tao, 2005). And moreover, they also cannot reflect and measure regional specialization directly. So, it is necessary to find a new way for letting these coefficients be easy and effective.

As we discussed above, location quotients can be used to measure industrial geographic concentration. Taking LQ as an example, if $LQ > 1$, it generally indicates a high concentration, and a higher LQ implies a higher concentration. Nevertheless, there exist a series of different situations for a certain industry. For instance, LQ in most regions may be greater than 1, but their variance may be very small or close to 1, or LQ in most regions may be less than 1, but their variance may be larger. Therefore, the specialization of an industry or a region is not only related to the value of LQ or FLQ or $FLQ^{\star}$ but also related to their variance (or discrete degree): If the variation of their value for a certain industry is very big, it indicates that this industry mainly localized in few regions, and its spatial distribution is more uneven, whereas if the variation is very small, it indicates its spatial distribution is more uniform; For a given region, if the variation of their value is very big, it indicates that this region mainly specialized in few industries. Consequently, we think the dispersion of location quotients value can reflect and measure the specialization of a region and the localization of a industry.

Mathematically, we can use the coefficient of variation (CV for short) to interpret the dispersion degree of a set of data. For non-negative series $X = (x_1, x_2, x_3, \cdots, x_n)$, the coefficient of variation is $CV_x = SD_x/\bar{X}$, where SD_x is the standard deviation of X, and $\bar{X}$ is the mean. CV can take 0 as the lowest value indicating complete uniformity while all x are equal (Note that it doesn't make sense if all x are equal to 0), and reach its upper bound $\sqrt{n-1}$ indicating absolute inequality when all x but one are equal to zero (Martin and Gray, 1971). We can define that it is of strong variability when $CV > 1$ in that standard deviation of the set exceed its mean, and on the contrary, it is of weak variability when $CV < 1$. Then, the regional specialization coefficient can be represented by a set of equations.

Coefficient for industrial localization:

$$\ell_i = \sqrt{\frac{1}{M}\sum_{r=1}^{M}\left(FLQ_{ri}^{\star} - \frac{1}{M}\sum_{r=1}^{M}FLQ_{ri}^{\star}\right)^2} \Bigg/ \frac{1}{M}\sum_{r=1}^{M}FLQ_{ri}^{\star} \tag{13}$$

Coefficient for regional specialization:

$$\ell_r = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(FLQ_{ri}^{\star} - \frac{1}{N}\sum_{i=1}^{N}FLQ_{ri}^{\star}\right)^2} \Bigg/ \frac{1}{N}\sum_{i=1}^{N}FLQ_{ri}^{\star} \tag{14}$$

According to properties of CV, ℓ_i takes 0 as the lower bound while industry i distributes uniform in all regions and all regions have an equal regional size. ℓ_i reaches $\sqrt{M-1}$ as the upper bound while it distributes only in one region. Similarly, ℓ_r takes 0 while region r has all industries and every industry has an equal E share in total E of corresponding industry in entire country, and it takes $\sqrt{N-1}$ while region r only specializes in one industry. In terms of the same way, we also can calculate CV of LQ and FLQ, and named CV_{LQ} and CV_{FLQ} respectively in order to carry out necessary comparative analysis. An obvious advantage of these coefficients are that they are easy to compute and can also be computed even if there exist some missing values or different regional and industrial classifications.

As a result, conclusions from the new method reveal some differences from those based on other methods. The regional specialization ranking comparative analysis between the new method and others (table 2) shows that the ranking of some regions whose size or industrial scale is less than others, such as Tibet, Qinghai and so on, dramatically decreased. Taking Tibet as an example, its regional specialization in 2003 ranked first in all coefficients except CV_{FLQ} and ℓ_r, the reason is that the proportion of their gross industrial output value or some industrial output value account for national total value was much less than that of others regions. However, these studies have obtained some common conclusions, for example, in terms of specialization, regions in Eastern China rank higher than those in central and western China (details will be shown in part four).

Table 2: Comparison of Regional Specialization Ranking Based on Different Coefficients

Region	ℓ_r (2003)	CV_{LQ} (2003)	CV_{FLQ} (2003)	Hoover (2003)	β_r (2003)	K-Spec (2003)	FR_r (2004)
Yunnan	1	3	1	3	2	6	6
Shanxi	2	4	3	7	5	5	3
Xinjiang	3	5	4	6	3	2	4
Heilongjiang	4	6	5	13	6	7	7
Tibet	5	1	2	1	1	1	1
Qinghai	6	2	6	2	4	4	2
Jilin	7	12	9	9	30	3	5
Guizhou	8	7	7	10	11	17	8
Hebei	9	15	10	19	9	22	17
Ningxia	10	8	8	5	13	13	10
Chongqing	11	14	12	14	10	10	——
Shannxi	12	13	13	15	17	18	14
Gansu	13	11	11	11	28	12	11
Inner Mongolia	14	10	14	8	7	11	9
Hainan	15	9	15	4	26	8	——
Guangxi	16	16	16	12	8	21	12
Guangdong	17	17	17	28	14	9	16
Sichuan	18	30	23	18	16	27	24
Liaoning	19	25	21	22	12	23	21
Hunan	20	23	18	21	23	28	23
Henan	21	18	19	20	18	26	15
Jiangxi	22	20	22	16	19	31	18
Zhejiang	23	21	20	26	15	14	20
Beijing	24	22	24	17	31	20	19
Shanghai	25	19	25	29	27	15	26
Jiangsu	26	26	27	31	24	16	29
Tianjin	27	24	26	24	20	19	22
Anhui	28	27	28	23	29	30	27
Fujian	29	28	29	25	22	24	25
Shandong	30	30	30	30	21	25	28
Hubei	31	31	31	27	25	29	13

Notes: FR_r coefficient is a coefficient created by Fan (2007).
Source: Hoover and β_r Coefficient calculated by Lu and Tao (2005); K-Spec (Krugman Specialization) Coefficient calculated by Guo and Yao (2007); F Coefficient calculated by Fan (2007).

Anyway, the specialization of a region or industry can be mathematically captured by the dispersion of LQ, FLQ and $FLQ^{\star}$. For an industry, if the CV of LQ, FLQ, and $FLQ^{\star}$ in all regions is high, it indicates this industry is more concentrated in several regions, so its localization level is higher. For a region, the CV of all industrial LQ, FLQ and $FLQ^{\star}$ is greater, it implies that this region mainly focuses on fewer industries, thus the specialization of this region is higher.

4. Measurement of Regional Specialization in China

The following measures of China's regional specialization from 1987 to 2007 are based on the method elaborated above with the sensitivity parameter $\delta = 0.3$. All the data come from the official statistical data (*China Industry Economy Statistical Yearbook* from 1988 to 2008), this paper picked the Gross Industrial Output Value of 26 two-digit industries of 31 provincial administrative regions as economic activity.

Firstly, we get a three-dimension database based on the computation of LQ, FLQ and $FLQ^{\star}$ of every industry and region. Then, we compute the CV of these coefficients by region and industry, namely CV_{LQ}, CV_{FLQ} and ℓ coefficients.

4.1. Overall Level of China's Regional Specialization

From the regional path perspective (See Table B1), there were 15 regions whose ℓ_r coefficients were greater than 1 in 1993, and the number of such regions increased to 24 in 2007. Compared with the situation in 1993, the regional specialization in 2007 of all regions has increased except for Fujian, Shanghai, Jiangxi, Gansu, Hainan, Guangxi and Tibet. Regional mean increased annually, the mean in 1987 was only 1.1201, and it increased to 1.399 in 1997 as well as 1.5336 in 2007. These results indicate that divergence of industrial concentration within inner-region has expanded and most regions specialized in developing some competitive edge industries, and these situations have become more and more obvious.

From the industrial path perspective (See Table B2), there were 14 industries (26 industries in total) which ℓ_i were greater than 1 in 1993, and it reached the amount of 18 industries in 2007. Localization of five industries (Mining and Processing of Ferrous Metal Ores, Food Manufacture, Tobacco Manufacture, and Smelting and Pressing of Non-Ferrous Metals) decreased in 2007 comparatively to 1993, and industrial mean was 0.982 in 1987, 1.2076 in 1997, and it increased to 1.3478 in 2007. This fluctuation showed that the divergence of most industries' concentration intensified from 1987 to 2007 and the regional specialization of industries has improved.

Consequently, our findings revealed that the overall level of regional specialization in China between 1987 and 2007 has remarkably increased, this phenomenon has also been partially verified by other researchers who used different methods (e.g. Bai et al., 2004; Liang and Xu, 2005; Fan, 2007).

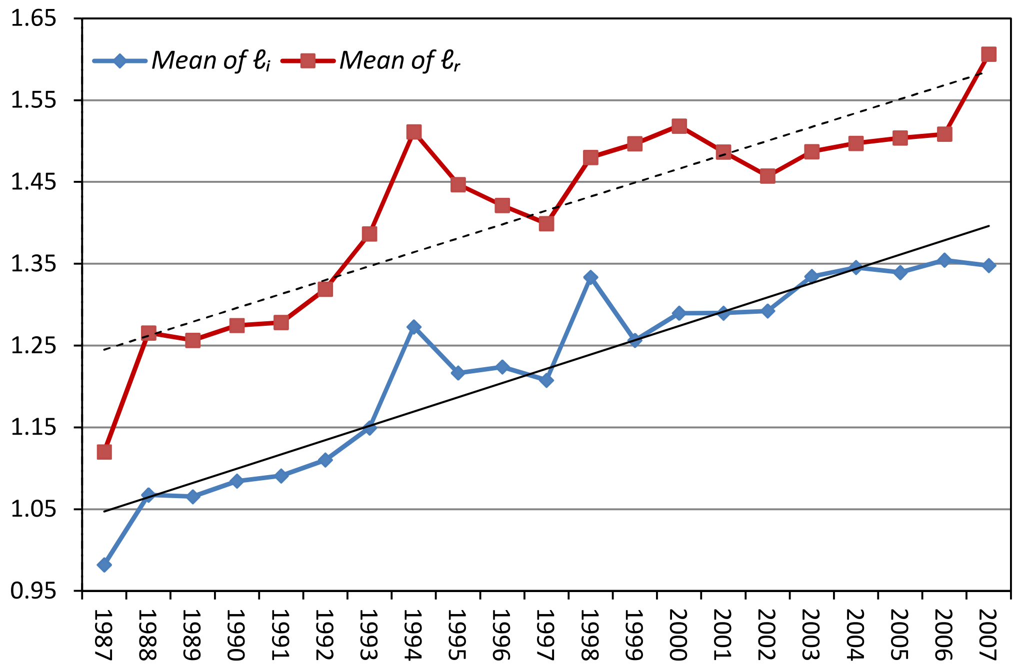


Figure 1: The Trend of Regional Specialization in China

4.2. Regional Differences of Specialization

There are clear regional differences in China's regional specialization. For instance, Western China's regional specialization is much greater than those of other regions, and Eastern China's regional specialization is the lowest. These conclusions can be supported by enough evidence in the whole period between 1987 and 2007. Taking the regional specialization in 2007 as an example, ℓ_r coefficient of all regions in Western China is greater than 1, but there are two-thirds of the regions in Central China whose ℓ_r are greater than 1, and in Eastern China, it is three-fifths. Similarly, the regional mean of ℓ_r indicates a much greater value in Western China (Table 3).

Table 3: Regional Differences of Specialization (2007)

	$\ell_r \geq 2$	$1 \leq \ell_r < 2$	$\ell_r < 1$	Regional Mean
Eastern China		Hainan(7)** Hebei(9) Tianjin(16) Zhejiang(17) Beijing(19) Guangdong(23)	Shanghai(25)** Fujian(28)** Jiangsu(29) Shandong(31)	1.1265
Central China	Shanxi(3)	Jiangxi(18)** Hunan(20) Henan(21)	Hubei(27) Anhui(30)	1.3031
Western China	Xinjiang(1) Qinghai(2) Yunnan(4)	Shaanxi(8) Gansu(10)** Chongqing(11)		1.9896

	Tibet(5)**	Guizhou(12)		
		Ningxia(13)		
		Inner Mongolia(14)		
		Sichuan(22)		
		Guangxi(24)**		
Northeast China	Heilongjiang(6)	Jilin(15)	Liaoning(26)	1.7343

Notes: the number in parentheses is regional ranking of ℓ_r; ** indicates ℓ_r in 2007 decreased comparison with in 1993; The regional economy pattern of mainland China has formed 4 major economic plates, Eastern China includes 10 provinces, Central China includes 6 provinces, Western China includes 12 provinces and Northeast China includes 3 provinces.

GIS map can intuitively reflects the regional differences of regional specialization between western China and eastern China (Figure 2). Thus all regions can be divided into two different groups by median after sorting ℓ_r in descending order (the median region is Tianjin whose ℓ_r approximately equals to 1.24). The higher specialization group include all regions in western China except Sichuan and Guangxi province, and the lower specialization group include all regions in eastern China except Hebei and Hainan province.

Figure 2: The GIS Map of ℓ_r Coefficient (2007)

The main reason of above regional differences is that most industries in China mainly locate in eastern areas and their scale are large, and western areas mainly focus on few industries with endowment resources. Number of industries whose output share rank in top 6 reveal the uneven distribution of industries (table 4), Shandong, Guangdong, Jiangsu, Zhejiang and Shanghai have more than 10 industries whose output share rank in top 6, especially, Shandong, Guangdong and Jiangsu province only have very few industries whose output share are not in top 6, industrial diversification leads to a lower specialization in these eastern coastal regions.

Table 4: Number of Industries whose output share rank in top 6

Eastern China			Central China		
Region	No. of Industries in top 6		Region	No. of Industries in top 6	
	1993	2007		1993	2007
Shandong	22	25	Henan	7	12
Guangdong	20	22	Hunan	2	5
Jiangsu	20	21	Hubei	5	3
Zhejiang	15	15	Shanxi	2	3
Shanghai	16	10	Anhui	1	2
Fujian	2	5	Jiangxi	1	2
Hebei	5	5			
Tianjin	0	3			
Beijing	3	2			
Western China			**Northeast China**		
Region	No. of Industries in top 6		Region	No. of Industries in top 6	
	1993	2007		1993	2007
Inner Mongolia	0	5	Liaoning	16	7
Sichuan	9	3	Heilongjiang	3	1
Yunnan	2	2	Jilin	1	1
Xinjiang	1	1			
Shaanxi	0	1			
Guangxi	2	0			
Gansu	1	0			

Notes: For every industry, we sorted in descending order and chose top 20% regions (approximate to 6 regions); Total number of industries are 26; Regions which do not show in this table indicate that they have no industry whose output share rank in top 6.

4.3. Industrial Differences of Specialization

Industrial differences of specialization were also obvious, a finding which can be confirmed by results from industrial localization coefficient ℓ_i of 2007 displayed in table 5. There were 4 industries whose ℓ_i coefficients were greater than 2 and 18 industries' ℓ_i were greater than 1. This situation is substantially different from that of past several years, there are 14 industries whose ℓ_i was greater than 1 in 1993, and only five industries whose localization coefficient decreased from 1993 to 2007. Localization level of most industries increased in the past two decades, this fact indicated that regional specialization in China improved exactly in this period.

Table 5: Industrial Differences of Specialization (2007)

$\ell_i \geq 2$	$1 \leq \ell_i < 2$	$\ell_i < 1$

[1]Tobacco Manufacture **	[5]Manufacture of Communication Equipment, Computers and Other Electronic Equipment	[19]Food Manufacture **
[2]Chemical Fiber Manufacture	[6]Mining and Processing of Non-Ferrous Metal Ores	[20]Mining and Processing of Nonmetal Ores
[3]Petroleum and Natural Gas Extraction**	[7]Manufacture of Textile Apparel, Footware and Caps	[21]Food Processing from Agricultural Products
[4]Coal Mining and Washing	[8] Textile Manufacture	[22]Beverage Manufacture
	[9]Mining and Processing of Ferrous Metal Ores**	[23]Manufacture of Special Purpose Machinery
	[10]Manufacture of Measuring Instruments and Machinery or Cultural Activity and Office Work	[24]Manufacture of Non-Metallic Mineral Products
	[11] Manufacture of Transport Equipment	[25] Medicine Manufacture
	[12] Electrical Machinery and Equipment Manufacture	[26]Raw Chemical Materials and Chemical Products Manufacture
	[13]Processing of Petroleum, Coking, Processing of Nuclear Fuel	
	[14] Manufacture of Metal Products	
	[15]Smelting and Pressing of Non-Ferrous Metals **	
	[16]Manufacture of General Purpose Machinery	
	[17] Paper and Paper Products Manufacture	
	[18]Smelting and Pressing of Ferrous Metals	

Notes: [ ] is the industrial ranking of ℓ_i; ** indicates ℓ_i in 2007 decreased comparison with in 1993, it implied the regional specialization of this industry reduced.

Moreover, some industries such as Food Manufacture, Mining and Processing of Nonmetal Ores, Beverage Manufacture, Manufacture of Special Purpose Machinery, and Medicine Manufacture which had very low localization level and their geographic distribution was much more uniform than that of others. Generally, special-resource-dependent industries, such as tobacco, gas, coal, metal ores and so on, exhibit higher localization level. These industries mainly concentrate in regions with rich resource, such as tobacco manufacture industry in Yunnan, petroleum and natural gas extraction industry in Xinjiang and Heilongjiang, coal mining and washing industry in Shanxi, mining and processing of non-ferrous metal ores industry in Henan, and so on; Industries with strong technical barriers, especially Manufacture of Communication Equipment, Computers and Other Electronic Equipment also had higher localization level. These industries are mainly distributed in regions with strong research and innovation ability such as Beijing, Shanghai, and Guangdong province.

5. Conclusions

This paper focuses on the measurement method of regional specialization. Firstly we improved FLQ and constructed $FLQ^{\star}$, and then a new coefficient of measuring regional specialization which are called ℓ coefficients (include ℓ_r and ℓ_i) has been put forward based on $FLQ^{\star}$ coefficient, and this method was used to measure a historical trend, regional and industrial differences of China's regional specialization. $FLQ^{\star}$ inherits advantages of LQ and FLQ coefficients, more importantly it considers the effect of regional scale of industry and can more accurately and effectively measure regional specialization and industrial localization. Unfortunately, ℓ coefficients cannot estimate regional specialization of a country by an integrated path, it must measure specialization in both industrial path and regional path. Moreover, there also exists a problem of the determination of parameter δ. However, it is an effective alternative method especially when complete and consistent statistical data are not available.

This paper also analyzed China's regional specialization empirically using ℓ coefficients and obtained conclusions similar to those obtained by other classical methods. These conclusions show that China's regional specialization increased from the late 1980s, in particular inland areas. Western China exhibits much higher regional specialization than do eastern coastal areas. The direct and main reason could be the uneven industrial geographic distribution: Most industries agglomerate in eastern developed areas, thus these regions do not specialize in few industries, but inland areas are still at a low level stage of industrialization and only focus on some resource-intensive and labor-intensive industries.

References

Amiti, M. (1998). New Trade Theories and Industrial Location in the EU: A Survey of Evidence. *The Oxford Review of Economic Policy*, 14, 45-53.

Bai, C., Du, Y., Tao, Z., and Tong, S. (2004). Local Protectionism and Regional Specialization: Evidence from China's Industries. *Journal of International Economics*, 63, 397 – 417.

Bedeian, A. G. and Mossholder, K. W. (2000). On the Use of the Coefficient of Variation as a Measurement of Diversity. *Organizational Research Methods*, 3, 285-297.

Brun, J.-F. and Renard, M.-F. (2002). International Trade and Regional Specialization in China. In Mary-Francoise Renard(Ed.) *China and Its Regions: Economic Growth and Reform in Chinese Provinces*, Cheltenham, Massachusetts: Edward Elgar Publishing Limited, 87-101.

Brülhart, M. (2001). Evolving geographical Concentration of European Manufacturing Industries. *Review of World Economics*, 137, 215-243.

Ellison, G. and Glaeser, E. L.(1994). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *NBER Working Paper*, No. 4840, http://www.nber.org/papers/w4840.

Ellison G. and Glaeser E. (1997). Geographic Concentration in US Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, 105, 889-927.

Flegg A. T. and Webber C. D. (1997). On the Appropriate Use of Location Quotients in Generating Regional Input-Output Tables: Reply. *Regional Studies*, 31, 795-805.

Flegg A. T. and Webber C.D. (2000). Regional Size, Regional Specialization and the FLQ Formula. *Regional Studies*, 34, 563-569.

Flegg, A. T. and Tohmo, T. (2010). Regional Input-output Tables and the FLQ Formula: A Case Study of Finland. *Regional Studies* (Submitted).

Fan, F. (2007). The Measurement of Regional Specialization. *Economic Research Journal*, 53, 71-83.

Guo, Z. and Yao, M. (2007). Industrial Agglomeration and the Estimation of Regional Specialization: 1994-2005. *Reform*, 19, 27-33.

Haggett, P. (1965) . Locational Analysis in Human Geography. London: Edward Arnold Ltd, 1965, 33-40.

Harigan, J. and Deng, H. (2010). China’s Local Comparative Advantage. In Robert C. Feenstra and Shang-Jin Wei(Ed.) *China’s Growing Role in World Trade*, University of Chicago Press, 109-133.

Holmes, T. J. and Stevens, J. J. (2004). Spatial Distribution of Economic Activities in North America. In J. V. Henderson and J.-F. Thisse(Ed.) *Handbook of Regional and Urban Economics*, volume 4, Amsterdam: North-Holland, 2297-2943.

Hoover E. M. (1936). The Measurement of Industrial Localization. *Review of Economics and Statistics, 18*, 162-171.

Krugman, P. (1991). Geography and Trade, Massachusetts, MIT Press.

Kim, S. (1995). Expansion of markets and the geographic distribution of economic activities: the Trends in U.S. *Quarterly Journal of Economics*, 110, 881-908.

Liang, Z. and Xu, L. (2005). Regional Specialization and Dynamic Pattern of Comparative Advantage: Evidence from China's Industries 1988-2001. *Review of Urban & Regional Development Studies*, 16, 231-244.

Lu, J. and Tao, Z. (2005). Regional Specialization and Inter-region Similarity: An Empirical Analysis of Economic Geography in China. *China Journal of Economics*, 1, 29-52.

Martin, J. D. and Gray, L. N. (1971). Measurement of Relative Variation: Sociological Examples. *American Sociological Review*, 36, 492-502.

McCann P. and Dewhurst J. H. L. (1998). Regional Size, Industrial Location and Input-Output Expenditure Coefficients. *Regional Studies*, 32, 435-444.

Naughton, B. (2003) . How much Can Regional Integration Do to Unify Chinese Markets?. In Nicholas Hope, Dennis Yang and Mu Yang Li (Ed.) *How Far Across the River? Chinese Policy Reform at the Millennium*, Stanford: Stanford University Press, 204-232.

Poncet, S. (2003). Measuring Chinese domestic and international integration. *China Economic Review*, 14, 1-22.

Young, A. (2000). The Razor's Edge: Distortions and Incremental Reform in the People's Republic of China. *Quarterly Journal of Economics*, 115, 1091-1135.

Appendix A: Proof for the Interval of CV, ℓ_i and ℓ_r Coefficients

A.1 Interval of Coefficient of Variation

For non-negative series $X = (x_1, x_2, x_3, \cdots, x_n)$, the coefficient of variation is $CV_x = SD_x/\bar{X}$, where SD_x is the standard deviation of X, and $\bar{X}$ is the mean. Further coefficient of variation is defined by,

$$CV_x = \frac{SD_x}{\bar{X}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(x_i - \frac{1}{n}\sum_{i=1}^{n}x_i\right)^2} \Bigg/ \frac{1}{n}\sum_{i=1}^{n}x_i$$

Obviously, if all elements of X are equal, i.e. $x_1 = x_2 = \cdots = x_n \neq 0$, then SD_x=0, thus $CV_x = 0$ takes the minimum. Following the statement of Martin and Gray (1971, equation (4)), if all elements but one are equal 0, i.e. for $i = h, h \subset [1, n]$, $x_h = k \neq 0$, otherwise for $i \neq h$, $x_i = 0$, then,

$$CV_{x,max} = \sqrt{\frac{1}{n}\left[\left(k - \frac{k}{n}\right)^2 + (n-1)\left(\frac{k}{n}\right)^2\right]} \Bigg/ \frac{k}{n} = \sqrt{\frac{1}{n}\left(k^2 - \frac{k^2}{n}\right)} \Bigg/ \frac{k}{n}$$

$$= k\sqrt{\frac{1}{n} - \frac{1}{n^2}} \Bigg/ \frac{k}{n} = n\sqrt{\frac{n-1}{n^2}} = \sqrt{n-1}$$

Thus the interval of CV_x is $[0, \sqrt{n-1}]$.

Now we repeat the assumptions in this paper. Assume that M is the total number of regions, N is the total number of industries; E_{ri} is the employment or output of industry i in region r; NE_i is the total employment or output of industry i in entire country, $NE_i = \sum_{r=1}^{M} E_{ri}$; TRE_r is total employment or output of region r, $TRE_r = \sum_{i=1}^{N} E_{ri}$; TNE is total employment or output of entire country, $TNE = \sum_{r=1}^{M}\sum_{i=1}^{M} E_{ri}$. Then,

$$FLQ_{ri}^{\star} = \frac{E_{ri}/TRE_r}{NE_i/TNE} \times \left[log_2\left(1 + \frac{E_{ri}}{NE_i}\right)\right]^{\delta} = \frac{E_{ri}}{NE_i} \times \frac{TNE}{TRE_r} \times \left[log_2\left(1 + \frac{E_{ri}}{NE_i}\right)\right]^{\delta}$$

A.2: Proof for the interval of ℓ_i

For any two regions $j, k \subset [1, M]$. If industry i distribute uniform in all regions, i.e. $E_{ji} = E_{ki} = a \neq 0$, and regional size of all regions are equal, i.e. $TRE_j = TRE_k = b \neq 0$, then,

$$FLQ_{ji}^{\star} = \frac{E_{ji}}{NE_i} \times \frac{TNE}{TRE_j} \times \left[log_2\left(1 + \frac{E_{ji}}{NE_i}\right)\right]^{\delta} = \frac{a}{Ma} \times \frac{Mb}{b} \times \left[log_2\left(1 + \frac{1}{M}\right)\right]^{\delta}$$

$$FLQ_{ki}^{\star} = \frac{E_{ki}}{NE_i} \times \frac{TNE}{TRE_k} \times \left[log_2\left(1 + \frac{E_{ki}}{NE_i}\right)\right]^{\delta} = \frac{a}{Ma} \times \frac{Mb}{b} \times \left[log_2\left(1 + \frac{1}{M}\right)\right]^{\delta} = FLQ_{ji}^{\star}$$

In terms of properties of coefficient of variation(CV), CV takes minimum 0 when all terms are equal, thus $\ell_{i,min} = 0$.

If industry i only localize in one region j, i.e. $E_{ji} = a \neq 0$ and $E_{ri} = 0, r \neq j$, then,

$$FLQ^{\star}_{ji} = \frac{E_{ji}}{NE_i} \times \frac{TNE}{TRE_j} \times \left[log_2 \left(1 + \frac{E_{ji}}{NE_i} \right) \right]^{\delta} = \frac{a}{a} \times \frac{TNE}{TRE_j} \times \left[log_2 \left(1 + \frac{a}{a} \right) \right]^{\delta} \neq 0$$

$$FLQ^{\star}_{ri,r\neq j} = \frac{E_{ri}}{NE_i} \times \frac{TNE}{TRE_r} \times \left[log_2 \left(1 + \frac{E_{ri}}{NE_i} \right) \right]^{\delta} = \frac{0}{a} \times \frac{TNE}{TRE_j} \times \left[log_2 \left(1 + \frac{0}{a} \right) \right]^{\delta} = 0$$

In terms of properties of coefficient of variation(CV), CV takes maximum when all terms but one are equal to zero, thus $\ell_{i,max} = \sqrt{M-1}$.

A.3: Proof for the interval of ℓ_r

For any two industries $m, n \subset [1, N]$. If region r have all industries and share of every industry are equal, i.e. $E_{rm}/NE_m = E_{rn}/NE_n = c \neq 0$, then,

$$FLQ^{\star}_{rm} = \frac{E_{rm}}{NE_m} \times \frac{TNE}{TRE_r} \times \left[log_2 \left(1 + \frac{E_{rm}}{NE_m} \right) \right]^{\delta} = c \times \frac{TNE}{TRE_r} \times \left[log_2 (1+c) \right]^{\delta}$$

$$FLQ^{\star}_{rn} = \frac{E_{rn}}{NE_n} \times \frac{TNE}{TRE_r} \times \left[log_2 \left(1 + \frac{E_{rn}}{NE_n} \right) \right]^{\delta} = c \times \frac{TNE}{TRE_r} \times \left[log_2 (1+c) \right]^{\delta} = FLQ^{\star}_{rm}$$

In terms of properties of coefficient of variation(CV), CV takes minimum 0 when all terms are equal, thus $\ell_{r,min} = 0$.

If region r only specializes in one industry m, i.e. $E_{rm} = d \neq 0$, and $E_{ri} = 0, i \neq m$, then,

$$FLQ^{\star}_{rm} = \frac{E_{rm}}{NE_m} \times \frac{TNE}{TRE_r} \times \left[log_2 \left(1 + \frac{E_{rm}}{NE_m} \right) \right]^{\delta} = \frac{d}{NE_m} \times \frac{TNE}{TRE_r} \times \left[log_2 \left(1 + \frac{d}{NE_m} \right) \right]^{\delta} \neq 0$$

$$FLQ^{\star}_{ri,i\neq m} = \frac{E_{ri}}{NE_i} \times \frac{TNE}{TRE_r} \times \left[log_2 \left(1 + \frac{E_{ri}}{NE_i} \right) \right]^{\delta} = \frac{0}{NE_i} \times \frac{TNE}{TRE_r} \times \left[log_2 \left(1 + \frac{0}{NE_i} \right) \right]^{\delta} = 0$$

In terms of properties of coefficient of variation(CV), CV takes maximum when all terms but one are equal to zero, thus $\ell_{r,max} = \sqrt{N-1}$.

Appendix B

Table B1: the Time Series of China's Provincial Regional Specialization (ℓ_r, 1987-2007)

Region	1987	1990	1993	1997	2000	2001	2002	2003	2004	2005	2006	2007
Beijing	0.7965	0.7769	0.8776	1.0270	1.2996	1.2409	1.0414	0.8888	0.8676	1.0100	1.0641	1.1081
Tianjin	0.5773	0.6131	0.7152	0.9569	0.8851	0.8465	0.8570	0.8211	0.8894	0.9905	1.1288	1.2439
Hebei	1.3547	1.3986	1.4530	1.4557	1.7938	1.8674	1.9104	1.8538	1.8319	1.7929	1.8087	1.7659
Shanxi	2.4469	2.4758	2.5384	2.7680	2.8403	2.8581	2.7917	2.7751	2.7862	2.8638	3.0062	2.9915
Inner Mongolia	1.1387	0.9747	1.0072	1.1650	1.3659	1.3294	1.3277	1.3600	1.4186	1.4132	1.4543	1.4674
Liaoning	0.7702	0.6828	0.7390	0.8532	0.9212	0.9715	1.0143	1.0601	1.0234	0.9682	0.9321	0.8894
Jilin	0.8286	0.8136	1.0269	1.3822	1.6462	1.8122	1.9181	1.9893	1.8308	1.6246	1.5176	1.4151
Heilongjiang	1.9831	2.3115	2.3615	2.3874	2.8011	2.8394	2.8153	2.7010	2.7479	2.7986	2.8627	2.8986
Shanghai	0.8506	0.8089	0.9220	0.9016	0.8837	0.7494	0.7567	0.8470	0.7984	0.8451	0.8621	0.9067
Jiangsu	0.6567	0.6999	0.7453	0.7282	0.7167	0.7488	0.7436	0.8211	0.7317	0.7807	0.8008	0.8104
Zhejiang	0.6969	0.6666	0.8049	0.7498	0.8109	0.8881	0.9114	0.9797	1.1331	1.1439	1.1884	1.1913
Anhui	0.6243	0.6373	0.6552	0.9338	0.9660	0.9116	0.8635	0.7833	0.7960	0.7921	0.7770	0.7079
Fujian	0.9263	0.7691	0.8613	0.7354	0.7123	0.7191	0.7503	0.7394	0.7733	0.7934	0.8128	0.8266
Jiangxi	0.4654	0.4453	1.4475	1.4883	0.8519	0.9734	0.9305	0.9892	1.1933	1.0389	1.0963	1.1811
Shandong	0.9812	0.5982	0.4943	0.4906	0.6422	0.6434	0.6604	0.6977	0.6304	0.5841	0.5672	0.5497
Henan	0.6878	0.6426	0.6566	0.7885	1.0079	1.0252	1.0733	1.0473	1.0640	1.0606	1.0356	1.0364
Hubei	0.8560	0.6963	0.7199	0.6064	0.7487	0.7520	0.7731	0.6626	0.8625	0.8532	0.8676	0.8416
Hunan	0.6700	0.8321	0.8414	1.0237	1.1390	1.1011	1.0259	1.0565	1.1146	1.1041	1.1098	1.0600
Guangdong	0.7153	0.6292	0.7562	0.9207	0.9233	0.9716	1.0203	1.0937	1.0481	1.0626	1.0427	1.0224
Guangxi	1.2105	0.9365	1.1487	1.1471	1.8755	1.5902	1.1773	1.1735	1.0469	0.9844	1.0168	1.0096
Hainan	——	3.5647	2.3947	1.8917	1.8051	1.5404	1.4433	1.2891	1.7137	1.7150	1.3271	1.8396
Chongqing	——	——	——	0.9518	1.5027	1.5184	1.4097	1.6974	1.5446	1.5490	1.6589	1.6985
Sichuan	0.6008	0.5310	0.4630	0.7742	1.0164	1.1103	1.0961	1.0843	0.9518	1.1011	1.0897	1.0274
Guizhou	0.7856	1.6744	1.5621	1.8643	1.9340	1.8843	1.8073	1.9531	2.0320	1.6598	1.6310	1.6714
Yunnan	1.0829	2.9550	2.8991	3.3289	3.3780	3.3619	3.3162	3.4781	3.1763	3.1446	2.8633	2.9260
Tibet	3.8031	3.7992	4.5537	2.4995	3.0632	2.6125	2.3415	2.3879	2.4065	2.1366	2.2325	2.7755
Shaanxi	0.8506	0.9193	0.9105	1.0346	1.2474	1.3421	1.4714	1.5629	1.5820	1.7219	1.6049	1.7692
Gansu	1.0427	0.8746	1.8576	1.6025	1.5868	1.5392	1.4793	1.5331	1.6738	1.6744	1.6886	1.7311
Qinghai	1.1181	1.5801	1.6088	2.1274	2.2448	2.1906	2.1619	2.1937	2.1971	2.6885	2.9114	3.0247
Ningxia	1.6952	1.3203	1.4718	1.4520	1.6243	1.5980	1.7463	1.8380	1.5087	1.5136	1.5959	1.5705
Xinjiang	2.2670	2.6102	3.1017	3.3326	2.8406	2.5536	2.5385	2.7451	3.0432	3.2114	3.2083	3.2033
Mean	**1.1201**	**1.2746**	**1.3865**	**1.3990**	**1.5185**	**1.4868**	**1.4572**	**1.4872**	**1.4973**	**1.5039**	**1.5085**	**1.5536**

Table B2: China's Industrial Localization (ℓ_i, 1987-1992)

Industries	1987	1992	Industries	1993	1997	2000	2003	2007
Coal Mining and Washing	1.3884	1.9369	Coal Mining and Washing	1.8831	1.8938	2.0848	2.1233	2.1315
Petroleum and Natural Gas Extraction	2.304	2.6375	Petroleum and Natural Gas Extraction	2.4618	2.5481	2.2865	2.1595	2.3019
Mining and Processing of Ferrous Metal Ores	1.7583	2.2808	Mining and Processing of Ferrous Metal Ores	2.0467	1.5647	2.2589	1.8147	1.5301
Building Materials, Mining and Processing of Other Nonmetal Ores	——	0.7091	Mining and Processing of Non-Ferrous Metal Ores	1.3024	1.2824	1.4811	1.2289	1.6899
			Mining and Processing of Nonmetal Ores	0.8302	1.1636	1.0151	1.2882	0.951
			Food Processing from Agricultural Products	0.6468	0.6874	0.916	0.9819	0.9363
Foods Manufacture		0.6297	Food Manufacture	1.1745	0.7383	0.7591	1.0098	0.9834
Beverage Manufacture	0.5894	0.6535	Beverage Manufacture	0.6976	0.8516	0.8888	0.9593	0.9348
Tobacco Manufacture		2.3702	Tobacco Manufacture	2.852	3.1029	2.8953	2.8603	2.685
Textile Industry	0.7811	0.9677	Textile Manufacture	1.0985	1.0165	1.2443	1.4014	1.5359
Sewing Industry	——	1.0678	Manufacture of Textile Apparel, Footware and Caps	1.272	1.2795	1.5794	1.6889	1.6589
Paper and Paper Products Manufacture	0.6487	0.6151	Paper and Paper Products Manufacture	0.628	0.6154	0.8517	0.9841	1.0619
Processing of Petroleum, Coking, Coal gas and Coal Products	1.4105	1.2166	Processing of Petroleum, Coking, Processing of Nuclear Fuel	1.2108	1.237	1.1422	1.2718	1.2224
Chemical Industry	0.5901	0.5341	Raw Chemical Materials and Chemical Products Manufacture	0.5197	0.5814	0.5801	0.528	0.5909
Pharmaceutical industry	0.6379	0.6145	Medicine Manufacture	0.5912	0.7392	0.6478	0.634	0.6032
Chemical Fiber Industry	1.4275	1.2617	Manufacture of Chemical Fibers	1.4962	1.4126	1.4186	1.9151	2.4088
Building Material and Other Nonmetal Ores Products	0.4274	0.4082	Manufacture of Non-Metallic Mineral Products	0.4433	0.469	0.5589	0.656	0.7585
Smelting and Pressing of Ferrous Metals	0.9478	0.9262	Smelting and Pressing of Ferrous Metals	0.9645	0.942	0.9141	1.0006	1.0243
			Smelting and Pressing of Non-Ferrous Metals	1.3608	1.3009	1.0826	1.0224	1.1063
Metal Products	0.5818	0.6495	Manufacture of Metal Products	0.6629	0.7424	1.0652	1.175	1.145
Mechanical Industry	0.5434	0.5977	Manufacture of General Purpose Machinery	0.6967	0.8114	1.027	1.0936	1.0839
			Manufacture of Special Purpose Machinery	0.6114	0.8301	1.0252	0.7792	0.7621
Manufacture of Transport Equipment	0.908	1.0946	Manufacture of Transport Equipment	1.0383	1.2639	1.4161	1.5027	1.3447
Electrical Machinery and Equipment Manufacture	0.7777	0.9436	Electrical Machinery and Equipment Manufacture	0.9222	1.0992	1.1592	1.3007	1.225
Electronic and Communication Equipment Manufacture	1.0247	1.3294	Manufacture of Communication Equipment, Computers and Other Electronic Equipment	1.4185	1.7841	1.7385	1.7026	1.8862
Manufacture of Instrument, Meter and Other Measuring Instruments	0.9297	0.9785	Manufacture of Measuring Instruments and Machinery or Cultural Activity and Office Work	1.0555	1.44	1.4927	1.6134	1.4817
Mean	**0.982**	**1.1101**	**Mean**	**1.1495**	**1.2076**	**1.2896**	**1.3344**	**1.3478**

Notes: The industry classification before 1993 is different, it is computed based on the original classification.