1988-2007 中国地区产业专业化的变迁：一个 U 形曲线

摘要： 本文旨在探究一个地区在经济增长过程中产业专业化或多样化的一般规则以及它的演化机制。我们也引入了基于回归的不平等分解方法，这时根据它们的重要性，按顺序排列对专业化产生影响的因素，以便提供解决各省间区域产业专业化不平衡发展的政策指导。通过面板数据的经验研究，本文发现在所有水平上（中国东部、中部、西部和省际水平）的地区产业专业化的演化出示了 U 形定律。这个定律的产生原因是消费需求的多样性，技术进步和交易成本递减协调作用的结果。我们对每种要素的显著性进行分析，发现地区产业专业化、工资、交通基础设施、对外贸易和外国直接投资（重要性程度依次下降）的初始水平是导致区域产业专业化不平衡发展的关键因素。最后，我们得出结论并提出了合理的政策建议。

关键词： 地区产业专业化，技术进步，交易成本

中国分类号：F062.9　　文献标识码：A

Changes in Chinese Regional Industry Specialization during 1988 and 2007: a U-shaped Curve

Jianhua Zhang,Wen Cheng

(School of Economics, Huazhong University of Science and Technology, Hubei, China)

Abstract: This paper aims to probe into the general rule of industry specialization or diversification in a region during its process of economic growth, and explore the mechanism of its evolution. We also bring in the method of regression-based inequality decomposition, and then sequence the factors, which have effect on specialization, according to their significance, so as to provide policy guidelines for solutions about the unbalanced development of regional industry specialization between provinces. Through empirical research on panel data, the paper discovers that the evolution of regional industry specialization at all levels (Eastern, Central, Western and the provincial level of China) shows the U-shaped curve law. The reason of this law lies in that diversity of consumers' demand, progress of technology and decreasing of transaction cost take effects coordinately. We conduct an analysis on the significance of each factors, then discover that the original level of regional industry specialization, wage, transportation infrastructure, foreign trade and foreign direct investment (FDI) , whose weightiness of importance decreases in turn, are the critical factors leading to unbalanced development of regional industry specialization. Finally, we draw the conclusion and propose correspondent policy suggestions.

Key Words: regional industry specialization; technological progress; transaction cost

1. Introduction

China has witnessed great boost over the past 30 years since Reform and Opening-up, during which regional industry structure has been adjusted and optimized gradually. While the coordinated growth of regional economy and Chinese Road To New Industrialization has been emphasized, regional industrial restructuring abstracts widespread concerning for the sake of its great significance on development strategy in China. Nevertheless, domestic scholars have long paid more attention to industrial structure upgrade, while the issue on quantity optimizing has seldom been involved, that is whether regional industry restructure should be realized by the way of specialization or that of diversification. On the background of the staggered transfer of industry in the eastern part of China, correspondent solutions are required urgently to issues as follows: how to obtain a correct understanding the evolution law of regional industrial specialization, how to adjust industry policy in all regions of China, how to promote the upgrading and optimizing of the regional industrial structure.

The eastern part of China is the earliest beneficiary of participating international industrial division. By undertaking industry transfer from developed industrial nations, the eastern part of China has found a well-grounded industry system, with the increasing level of regional industry diversification. However, attributing to multiple factors such as increasing of labor cost and raw materials, appreciation of the RMB, and the cancel of tax reimbursement for export, many labor intensive industries in eastern part of China have become to lose their comparative advantage, leading to the downturn trend in economic efficiency. Therefore, eastern coastal areas took the lead in implementing the policy of industry transfer known as “emptying the cage, removing the bird,” focusing on high value-added industries such as advanced manufacture and modern service industry, meanwhile transferring out labor intensive industries which are of energy-hungry, heavy pollution and low value-added. Owing to what have been discussed above, the level of regional industry specialization is increasing step by step in those areas. But for the sake of a vast territory of China, quite big differences of economic development level between areas probably not lead to the same problems and challenges. In the face of the distinguished regional disparity, how to adjust polices to local conditions and work out applicable polices so as to promote the optimizing and upgrade of regional industry structure, that is a crucial problem.

Actually, the development of specialization in one country or area goes along with the evolution of division. In recent years, by conducting researches on the history of regional industry structure’s evolution, Imbs and Wacziarg (2003) found that: in the early phase of industrialization, the industry structure tends to be diversified along with the progress of industry division; whereas when per capita income ascending to a high level, industry concentration become manifested, and industry structure gets specialized again, thus what illustrated above help them to draw an U curve of the index of regional industry specialization. They even predict that it’s a general law .And they cite China as the evidence of low income country, then illustrate a curve to display the Gini Coefficient’s dropping process of Chinese regional industry specialization from 1983 to 1988. Unexpectedly, the economy in China has enjoyed a surging, which is convinced by the growth of nominal GDP per capita from 1366 RMB in 1988 to 19524 RMB in 2007, and witnessed a huge leap of Chinese economy to a new era. Therefore, it's necessary to further discuss whether the evolution of regional industry specialization in China conforms to the law shown in developed countries. More importantly, it is well worth to discuss the mechanism of regional industry specialization, combined with the testimony of data from 1988-2007 in China. Besides, the following issues remain to be concerned: Whether the development of regional industry specialization between provinces show significant difference under the background of unbalanced development of China economy? Which factors lead to the difference mentioned above and which factor is dominant? How to wipe off the difference by policy adjustment? In a word, only if know well about the universality and particularity can we come up with the policy on the optimizing and upgrade of regional industry structure in an objective and reasonable way.

The paper is structured as follows: In section two, we take literature review on the empirical research, causal analysis, and latest development on regional industry specialization. In section three, we carry out the nonparametric fitting drawing of regional industry specialization with per capita GDP, covering provinces, areas and all the country level. And random effect model is applied to obtain the specific parameter of quadratic curve, then we conclude the general statistic law of regional industry specialization. In section four, we conceive a mathematical model to explore the evolutionary mechanism in a micro view. Then we introduce total factor productivity and transaction cost into Ricardian model which explains the law that regional industry specialization is positive correlated with total factor productivity, but negative correlated with transaction cost. Finally we use the dynamic panel data of eastern, middle and western parts of China to confirm the mechanism empirically. In section five, we use the regression-based inequality decomposition method to figure out each factors and their contribution weightiness, which result in unbalanced development of regional industry specialization. The six section of this paper draws the conclusion and proposes detailed policies of industry structural adjustment.

2. Literature review

Regional specialization refers to the fact that production factors in one area concentrated in some specific industries，as a result, minority of industries contribute to a large part of the total value of out-put in this area. So far, there are two ways of thinking for the research of regional industry specialization: one is the study on comparative specialization among various regions, named “industry specialization between areas”; another is to study the modulus specialization in one region itself, which is claimed as “regional industry specialization”. The former one means to what extent industry structure in one area is different from that of other area, which is used to explain phenomenon such as “regional industry structure convergence” and “enhancement of integration of regions”. The last one holds the view that specialization and diversification of industry structure in one region itself have nothing to do with that of other regions. What the author concerns is the latter one “regional industry specialization”.

As far as now, there are abundant empirical papers on the laws of the evolution of regional industry specialization. Kim (1995) finds that the time series data of regional industry specialization index and average industry concentration demonstrate the law of U-curve. Another scholar Fan (2004) uses interlocal comparative specialization index and regional specialization to deduce that: China is on the ascending part of the upside-down U Curve, and is during the phase of the regional integration's advancing. However, relatively rare papers, which theoretical and empirical methods are combined, are found to study evolutionary laws of regional industry specialization. What's more, on account of the different factors, the evolutionary law regional industry specialization presents is different from that of interlocal comparative specialization.

Through using the Gini Coefficient of regional industry specialization which is present by the percentage of employment and added-value in the different industries, Imbs and Wacziarg (2003) conduct a statistical research on the evolution of regional industry specialization, covering 30 years past and involving 50 countries' data. This research discovers that regional industry structure tends to be diversified in the initial development of economy, whereas the industry turns out to be specialized again when per capita income reach 9000-10000 dollars(constant dollars terms in 1985). So in this way, the Gini coefficient will present a U-shaped curve along with the growth of per capita income. The researches of domestic scholars support this view at large. For instance, He and Xie (2006) calculate the provincial Gini coefficient of China manufacture from 1980 to 2003, and then find the integral specialization of China weakening gradually, and the industry structure more diversified. Besides, Gini coefficient measured by employment dropped from 0.57 in 1980 to 0.52 in 2003. But the specialization of China Eastern coastal areas has witnessed advancement since the end of the 1990s. By using the data of industry added value to calculate Gini coefficient of regional specialization for each province from 1988 to 2002, Lin (2007) discovers that the Chinese industry specialization reveals a law of U-shaped curve, and Gini coefficient of added-value in the different industries went up slightly after it reached the bottom in 1997. She also finds from the provincial data that: the specialization level is the highest in the Western, lower in the East, and the lowest in the middle part of China. Further, she categorize the sequence of Gini coefficient into four sorts: descending, U-shaped, ascending, invariant. Shi and Zheng (2010) noticed that: the evolution of manufacture structure, tertiary industry structure and the whole industry structure all present the law of U-shaped curve as per capita income increases.

Nevertheless, majority of the empirical researches on regional industry specialization in China make use of integral or provincial indicators of specialization to depict the tendency year on year. But for the sake of vast territory of China, the huge developmental discrepancy between areas makes it difficult to find the common law among provinces, only by classifying and summarizing time series and the data of an individual province. In the text below, we will combine the regional industry specialization index with per capita GDP, then include all levels in a single and more comparable framework, so as to draw an universal statistical law of regional industry specialization evolution.

Concerning the explanation of U-shaped curve of regional industry specialization, traditional trade theory claims that the extension of comparative advantage is a vital cause of regional industry diversification. In those theories, Ricardian model emphasizes on the comparative differences of productivity, while Heckscher-Ohlin model highlights the talents of factors. Dornbusch, Fischer and Samuelon (1977) propose an extended Ricardian Model containing continuous commodity sets, in which the commodities' coverage of comparative advantage will be expanded along with the promotion of labor productivity. As a result, the industry structure in one country tends to be diversified, while the cutting-down of cost and tariff will enlarge the extent of tradable commodity. On the basis of analysis above, Imbs and Wacziarg (2003) regard transport cost as the impetus of regional industry specialization, and hold the view that when the driving force of diversification exceeds that of specialization, the industry tends to be diversified, conversely specialized. But they only make inference theoretically, not test it by empirical data. Mentioning the causes that lead to U-shaped distribution of Chinese regional industry specialization, Lin (2007) believes that: in large parts of provinces in west and some provinces in Central China, Gini coefficient of regional industry specialization is high because of its invariant industry structure; in Eastern China and majority of provinces in Middle China, Gini coefficient is relatively low in that industry structures in these areas are distributed evenly and pillar industry is rare; In a few provinces of Eastern China which are more advanced, Gini coefficient is relatively high for the reason that industry structures in these areas are distributed evenly and there is a pillar industry. Shi and Zhen (2010) both believe that the base of development of regional specialization is to foster some dominant industries, whose growth lie in internal differentiation and upgrades, all of which can not exist without the support of innovation. Huang and Shu (2007) hold the view that the openness degree of trade promotes advantaged industry of high productivity, as well as the advancement of regional industry specialization , hence force China to lay stress on sectors of high efficiency and drive economy's long-term growth. For the theories home and abroad, each all has their own merits, but none of which give a persuasive explanation of the mechanism of the U-shaped curve shown by regional industry specialization. Therefore, the aim of the paper is to conceive a micro basis which tries to explain the mechanism of the U-shaped curve.

Compared to the evolutionary law of U-shaped curve which is seldom touched upon, many scholars abroad, more often than not, make researches on whether specialization or diversification is helpful to the growth of regional industry (Glaeser, Kallal, Scheinkman and Shleifer, 1992), sufficient employment and stability (Malizia and Ke, 1993), and the innovation in these regions (Feldman and Audretsch, 1999). These researches are all conducted from the view of MAR Externality, Jacobs Externality or Porter Externality.

So far, a large number of domestic researches on regional industry specialization pay more attention to the critical review on its effects. For instance, Bo (2007) has carried out a research on regional industry specialization's restriction and promotion on industry growth. However, very few scholars conduct research on the function and weightiness of each factor, and how to apply policy to lead a balanced development of regional industry specialization.

On account of adjustment of economical industry and settings of developmental strategy, it becomes more important to research the impacting factors of the evolution of the regional industry specialization, as well as the weightiness of each factor. Fortunately, improvement on the unequally decomposition to the regression function (Wan, 2002) provides a available tool for it. The new method possesses many advantages (Wan, 2008): Firstly, there is no restriction to unequal indicators (Gini coefficient, Tell coefficient and all the other unequal indicators can be used). Secondly, all influencing factors are permitted in the method, so the result of decomposition is accurate and convincible. Thirdly, no equation is required, but no estimation of regression functions. Lastly, there is no restrain to regression function, which can be highly linear or include interacting items. Except for traditional income distribution, the decomposition technology is widely used to explain the reason why differences exist between financial developments of provinces (Li, Ran and Wan, 2007) and the decisive factors which lead to the disparity of innovation (Wan, Fan and Lu, 2010). But very few scholars do further research on the unbalanced development of regional industry specialization in China. Therefore in this paper, an unequal decomposition of regression function will be used to analyze the issue further in order to provide foundation for the policy of the regions.

3. Statistical interpretation on the evolution tendency of regional industry specialization

In this part, supported by China's industrial statistics yearbook and China Statistical Yearbook, we will calculate the indicators of regional industry specialization (expressed by employment and added value) by the calibers of industry and the whole sectors respectively. Later, nonparametric LOWESS (locally weighted scatter plot smoothing) is applied to fit the coefficient and actual per capita GDP. In this way, we can ascertain a U-shaped law directly. Finally, the form of function and the bottom of the U-shaped curve can be decided through fitting the parameter of quadratic curve.

3. 1. Selection of the measuring index and the explanation of LOWESS

Gini coefficient, Herfindahl index, and Variation Index can be made use of to measure the degree of specialization and diversification of the regional industry structure. Because factors of relative scale between areas are not involved when calculating the degree of specialization in one region (Fan, 2007), and the outputs of each measuring index are highly coordinated (Imbs and Wacziarg, 2003), the paper employ Gini coefficient to measure the degree of regional industry specialization (variables are from 0 to 1). The hypothesis is: the bigger the result is, the higher the degree of regional industry specialization is, the lower the degree of diversification is. We use a scatter distributed Gini coefficient formula proposed by Sen.[1]

$$Spec = 1 + \frac{1}{n} - \frac{2}{n^2 \mu} \sum_{i=1}^{n} (n+1-i) x_i \qquad (1)$$

Xi refers to the employment and added value of industry *i* in one area, and *Xi* is lined in an ascending sequence of $x_1 < x_2 < \cdots < x_{n-1} < x_n$. n represents the number of samples, $\mu = \frac{\sum_{i=1}^{n} x_i}{n}$ refers to average of samples.

The formula（1）can be simplified as:

$$Spec = \frac{n+1}{n} - \frac{2(n+1)}{n^2\mu}\sum_{i=1}^{n} x_i + \frac{2}{n^2\mu}\sum_{i=1}^{n} ix_i = \frac{2}{n}\frac{\sum_{i=1}^{n} ix_i}{\sum_{i=1}^{n} x_i} - \frac{n+1}{n} \quad (2)$$

The index of regional industry specialization Spec reflects the degree of specialization or diversification of the industrial structure in one area: the bigger the index is, the higher the degree of specialization in this area is; the smaller the index is, the higher the degree of diversification in this area is. We also refer to China Statistical Yearbook and Compilation of statistical documents (1949-2004) for the nominal GDP and CPI index at national and provincial level, and then calculate national and provincial actual per capita GDP (data in 1985 used as the calculating base). Further, LOWESS is applied to fitting the coefficient of the actual per capita GDP with the index of regional industry specialization.

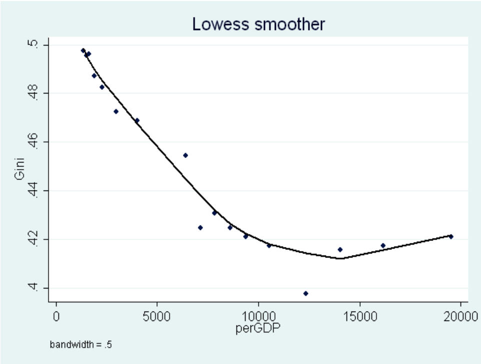
LOWESS is also known as locally weighted scatter plot smoothing, which is proposed by Cleveland (1979) . The core of the technology is described as: First, using the method of weighted AVG for each sample plot to select a subsidiary sample (the proportion of the subsidiary sample to the total sample is *f*) [2], then regression is applied to the subsidiary sample to obtain the smoothing output of *Yi*. In this way, the interrupts of the plot that is far from *Xi* is removed. Further, the regression in LOWESS is attached with weightiness, so the core plot（*Xi, Yi*） is of the highest weightiness, and the farer the plot of subsidiary sample is, the lower the weightiness is. Owing to the overlap of the subsidiary sample section, the smooth of the fitting results can be achieved.

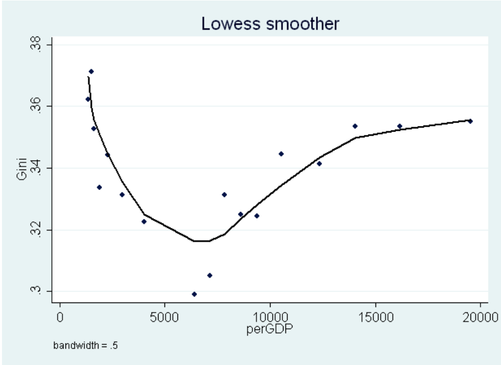
In following two sections, we will conduct research on the national industry specialization in an overall view, as well as the evolutionary law of the regional industry specialization of Eastern, Middle, Western part of China respectively.

3. 2. the U-shaped law of the national industry specialization

In the analysis of manufacture industry, we use the employment and added value data (classified by sectors) from 1988-2007 cited from China's industrial statistics yearbook, to calculate the integral regional industry specialization in nationwide (Gini),[3] and fit the annual nominal per capita GDP by LOWESS,[4] the results are described in the following two Figures:

Figure 1. Employment index to per capita GDP **Figure 2. Added value index to per capita GD**





We can directly find from the two Figures that: the relation between regional industry specializations of manufacture with nominal per capita GDP can be drawn in a parabola, of which the opening is upwards. That is to say, the coordination of the regional industry specialization with the development of economy can be illustrated in a U-shaped curve.

Thereafter, we will use simplified quadratic polynomial to work out each parameter, so as to deduce the income level when the U-curve touches the bottom.[5] In order to remove the impact of price, we use regional industry specialization index to do regression on actual GDP and its quadratic item, as well as nominal GDP and its quadratic item. By conducting a Unit Root Test of the residual of the regression (applied to actual GDP) and its square item, with regional industry specialization, we find that the residual shows stability. It declares that actual GDP and its square item is coefficient with the regional industry specialization index, which is also can be seen as a long-term stable relationship. For the sake of transnational compare, we will summarize the calculating results of 14 OECD countries and show them in the third and fifth columns of Figure 1. It can be concluded from the grids that: in both China and 4 OECD countries, the regional industry specializations show apparently the law of a U-shaped curve along with growth of economy. Making comparison between Column 2 and Column 3, as well as Column 2 and Column 3, it is found that: no matter the data of employment or that of added value, the turning point income of China is slightly below that of OECD countries, which can be explained by the catching-up strategies launched by Going-after countries. By making comparison between Column 2 and Column 3, as well as Column 2 and Column 3, we also find that: when data of employment is used to calculate and fit the U-curve, the turning point of it will lag behind that calculated by data of added value. This is probably because the labor is viscous when it transfers to higher value-added industry, thus lags behind the variation of industrial added value.[6]

Table1. The regression results of regional industry specialization index toactual per capita income and there quadratic items (China and OECD countries)

Variables	Employment of China （1988-2007）	Employment of OECD countries （1960-1993）	Industry added value of China （1988-2007）	Industry added value of OECD countries （1960-1993）
Actual per capita GDP	-0.0826*** （14.77）	-0.0300*** （11.67）	-0.0417*** (2.89)	-0.0101*** （3.82）
Quadratic term of GDP	0.0106*** （10.99）	0.0020*** （13.44）	0.0077*** (3.07)	0.0007*** （5.46）
Intercept	0.5677*** （84.00）	0.6590*** （47.64）	0.3815*** (21.87)	0.5029*** （35.55）

R-squared	0.9719	0.1370	0.4109	0.0750
Lowest point(actual)	3880	4369	2711	3490
Lowest point(nominal)	14246	16041	8913	11474

Notice: (1) Regression results are computed by STATA11.0. The results in the brackets are absolute value of T statistic, *** indicates the results have pass the significance inspection at the Significant Level of 1%. (2) Minimum of income that computed by data of OECD countries are converted into actual RMB income, according to purchasing power parity in Penn World Table.[7]

For the total industry analysis, the data we quote is the index of "the population of employment grouped by sorts of industries in each region" from the report in China Statistical Yearbook (1988-2002). The index covers all the 16 one-digit industries (from A to P) in The National Economy Industry Classification and Code.

Unfortunately, the Yearbook did not report the index since 2003, but shifted to the index of "the population of staff grouped by industries in each area", which covers 19 one-digit industries except international organizations (A-S, except for T). The range of the statistical calibers about the latter index is relatively narrowed. As a result, the absolute value of Gini Coefficient changes, but it will not affect its variation tendency. Therefore, we apply regression to the two columns of data respectively (Figure 2). The results show that: the level of regional industry specialization was dramatically declining along with the growth of actual per capita GDP from 1988 to 2002; whereas the level of regional industry specialization was advancing as actual per capita GDP was increasing. The turning-point of the U-shaped curve is at the GDP 11129 RMB. As for the index "Added value classified by industries", it was only reported from 2004, so we can not make an effective analysis on it.

Table 2. Regression results of China Industry Employment Index to actual per capita GDP by time slots

Variables	Industry employment data of China （1988-2002）	Industry employment data of China （2003-2007）
Actual per capita GDP	-0.0218*** （10.32）	0.0039*** （6.11）
Intercept	0.7785*** （204.04）	0.7182*** （275.05）
R-squared	0.8912	0.9257
Lowest point(actual)	3322	
Lowest point(nominal)	11129	

Notice：Regression results are computed by STATA11.0. The results in the brackets are absolute value of T statistic.

With regard to the triple transforms of industrial structures during the process of industrialization, it also contributed to the promotion and strengthening of the U-shaped curve law reflected by regional industry specialization. Because when a region is at the initial time of industrialization in which there is only agriculture, the level of regional industry specialization is the highest. While as the process of industrialization is accelerated, the sorts of industrial departments are increasing, and more labor force is flushing in, the regional industry specialization inclines to be diversified. Entering the later stage of industrialization, the rapid development of service industry attracts most part of economy entity and this industry gradually turns to be specialized.

3. 3. The U-shaped Curve law of regional industry specialization in each province, Eastern, Middle, and Western Part of China.

We use employment data which is aggregated by industries (China Industrial Economy Yearbook 1988-2007) to calculate the regional industry specialization of 28 provinces and districts, and LOWESS is applied to fitting the curve according to the ranges of Eastern, Central, Western and the overall country.[8] Then, we use the function of quadratic polynomial to work out the parameters. All of OLS, REM and FEM can be used to estimate the static panel data model. According to the result of Lagrange multiplier test, OLS is rejected. Hence, Hausman test is conducted and the method of REM is accepted. The results are shown in Figure 3.

Figure 3. The curve of regional industry specialization index to per capita GDP at all levels

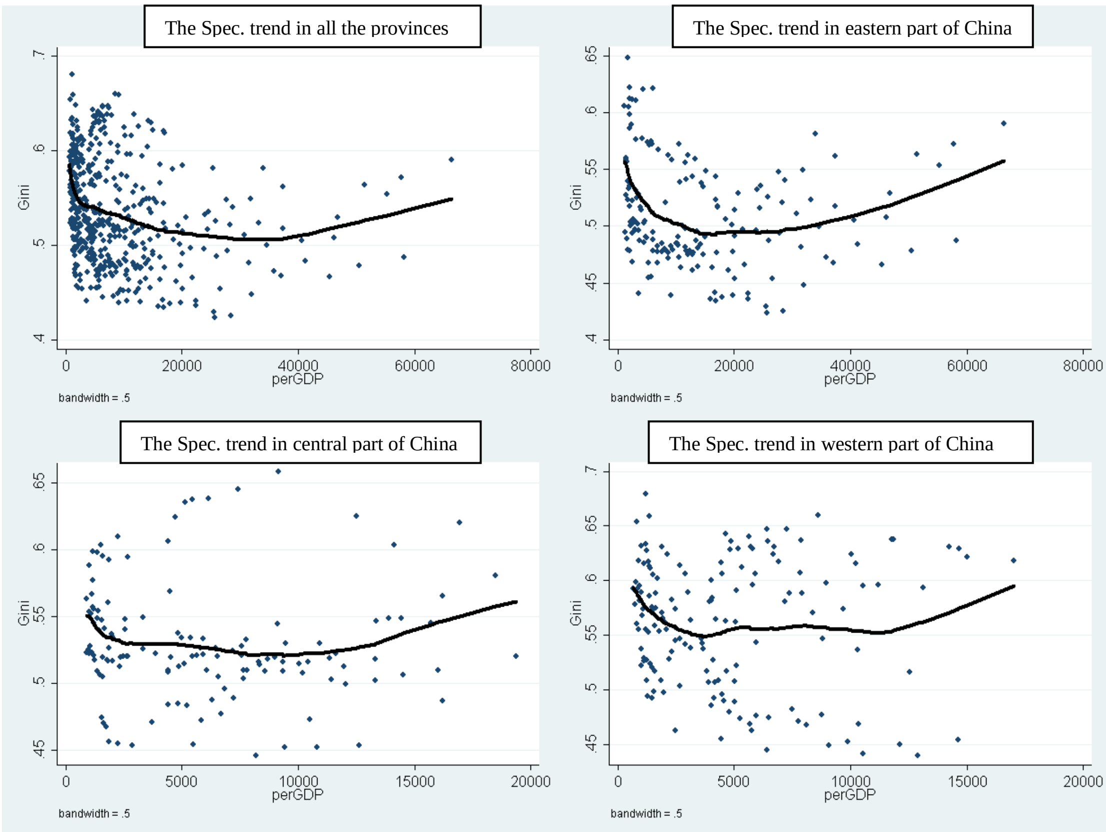


Table 3. the Regression of regional industrial specialization to actual per capita income and its quadratic items

Variables	All the provinces	Eastern part of	central part of	western part of

	（1988-2007）	China（1988-2007）	China (1988-2007)	China（1988-2007）
Actual per capita GDP	-0.0141***（7.51）	-0.0201***（8.96）	-0.0395***（5.68）	-0.0597***（4.09）
Quadratic item of GDP	0.0010***（6.08）	0.0014***（7.84）	0.0081***（5.76）	0.0013***（3.56）
Intercept	0.5597***（60.96）	0.5571***（41.02）	0.5658***（32.37）	0.5869***（32.25）
R-squared	0.1134	0.3441	0.2072	0.1219
Lowest point(actual)	6933	7065	2433	2296
Lowest point(nominal)	28687	29268	8290	8039
samples	476	170	136	153

Notice: Regression results are computed by STATA11.0. The results in the brackets are absolute value of T statistic, *、**、*** indicate the results have pass the significance inspection at the Significant Level of 1%. R-squared is within effect R-squared。

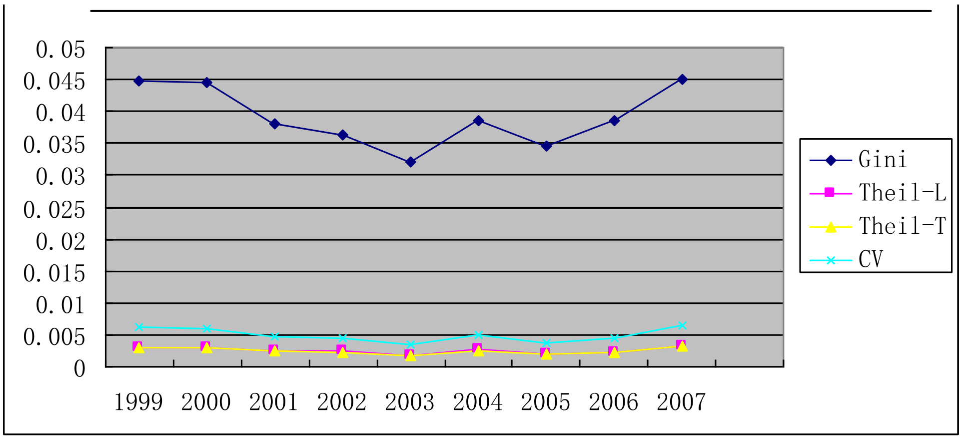
Combing the fitting results, we discover that the U-shaped curve law in East, Middle, and West Part of China is significant: The coefficient of actual per capita GDP and that of its quadratic item are testified to be significant at the level of 1%, and the R^2 are reasonable in the group. This signifies that: At the provincial level and the level of Eastern, Middle, Western Parts, the developments of industry structures all experience a course of diversification first and later specialization. Among all the provinces, the ones that experience the integral U-shaped curve are shown in the table below.

Table 4. The evolutionary phases of the provinces which experience Integral U-shaped curve (1088-2007)

Province	Beijing	Tianjin	Shanghai	Zhejiang	Jiangsu	Guangdong
Falling Phases	1988-2001	1988-1998	1988-1999	1988-2002	1988-2003	1988-1999
Rising Phases	2002-2007	1999-2007	2000-2007	2003-2007	2004-2007	2000-2007
Province	Fujian	Shandong	Heilongjiang	Jiling	Henan	Guizhou
Falling Phases	1988-1996	1988-1998	1988-1996	1988-1996	1988-1996	1988-1993
Rising Phases	1997-2007	1999-2007	1997-2007	1997-2007	1997-2007	1994-2007

According to the table, do differences exist between the provinces which are in the phase of specialization? We choose the period of 1999-2007 as the comparable phases of 12 provinces, hence sufficient samples can be included, besides, three quarters of the sampling provinces are gradually entering the phase of ascending, which is also acceptable in the research. In order to calculate the regional industrial specialization index of the 12 provinces (experiencing integral U-shaped curve), four unequal indicators are applied: Gini coefficient, Coefficient of Variation, and two generalized entropy coefficient Theil-L and Theil-T. Through the calculation above, the unequal evolutionary law (1999-2007) is shown in Figure 4. The Figure shows that Gini coefficient was keeping declining from 0.0447 in 1999 to the bottom, 0.0320 in 2003. Then after, Gini coefficient experienced constant rising and reached the top 0.0450 in 2007. The variation depicted above coincides with the U-shaped curve law. Besides, Coefficient of Variation, Theil-L and Theil-T display the same tendency of U shaped, only but the value is relatively small, and the variation is gentle. On account of all that discussed above, although all of the provinces are at the phase of ascending, it is a urgent problem to narrow down the discrepancy of the development of the regional industrial specialization between provinces.

Figure 4. Unequal tendency of regional industrial specialization in China



4. Factor Analysis on the Evolutionary Tendency of Regional Industrial Specialization

In this section, we will explore the factors that drive the evolution of regional industrial specialization: consumers' preference, the decision driven by maximized manufactures' profit, improvement of institution and how their interaction makes the U-shaped curve. We propose a micro theoretical model to explain the mechanism, and apply a macro statistical model to discuss the macro phenomenon that is lead by micro decisions. Finally, we use the panel data summarized in terms of provinces, East, Middle, and West (1988-2007) to conduct an empirical test.

4. 1. Micro mechanism of regional industrial specialization's evolution

The specialization means an economical entity distribute the majority or even all of its resources to one or serial economical activities for a long time. The basic forms of specialization are: department specialization, product specialization, and function specialization. The products specialization means an enterprise only produce and assemble the same kinds of products or the ones of the similar manufacturing process. Along with the development of products specialization, intra-industry at the medium will be brought in, and goes the specialized manufacturing. Besides, the department specialization in one area will promote further distribution and the concentration of industries. More over, function specialization means: The form of distribution experience the process from intra-product distribution to intra-industry distribution, and to distribution by the segments of the industrial chain. In this paper, we will stress the industrial specialization of one region at the department level, while the micro modeling foundation of industrial specialization happens to be products specialization. As a result, we proposed a simplified two-department model, and combine it with the view of products specialization to analyze the evolutionary mechanism of regional industrial specialization.

Why does technical progress facilitate diversification of products? On the base of converse tangent utility function, Cao (2008) found a statistical model to explain how technical progress facilitates diversification of products. But the model doesn't include the effect of the specialized distribution, brought by transaction costs' decrease, on the sorts of the products. On account of this aspect, the model can not explain the U-shaped curve that regional industrial structure tends to be diversified first, later to be specialized. We modify and extend the model, so as to explain the evolutionary mechanism of regional industrial specialization in a better way.

We assume that a region only has resource-dependent industries, like agriculture and excavation. The industries, which are of low degree of processing, produce relatively fewer kinds of products, leading to a single regional industrial structure and high level of regional industrial specialization. We presume the number of the sorts of products manufactured by resource-dependent industries is a Constant *m*, and its production function is:

$$Q_0 = A_0 L_0 \tag{3}$$

A_0 is the technology level of resource-dependent industries; L_0 is labor input of resource-dependent industries.

Supposing P_0 is the price of minerals and agricultural products, the income of labor in resource-dependent industries is:

$$w_0 = P_0 A_0 \tag{4}$$

As the technology in this region is progressing, the region gradually shifts to deep processing of resources products and production of consumption goods. The sorts of industry sectors are gradually completed and regional industrial structures tend to be diversified, but the level of specialization is lowered and the development of if enters the second phase.

In the third phase, for the sake of transaction cost's declining, the manufacture will abandon producing the products of low labor productivity, then shift to products of comparative advantage. In this stage, products of comparative advantage will gradually turns to be the predominant industry. Therefore, the level of regional industrial specialization will go up again. The latter two stages can be illustrated as follows:

The industrial sectors of Non resource-dependent enterprises produce n sorts of differentiated products, assuming the production function of industrial products is:

$$Q_i = A_i (L_i - f_i + T_i), i = 1, \cdots, n. \tag{5}$$

A_i represents the technical level of Product i, L_i is the amount of labor needed to produce Product i, f_i is the fixed input needed to produce Product i. T_i is the transactional cost incurred in purchasing Product i from other region, which helps to form trade barriers that protects diversified manufacture. When transaction cost is reduced by improvement of communicational, informational infrastructures, or institutions, manufacturers will give up parts of products of low labor productivity and shift to those of high labor productivity.

The monopoly manufacturers will determine the quantities and prices according to the principle of maximized profit, which can also be stated as Marginal Return Equals to Marginal Cost:

$$P_i\left(1-\frac{1}{\varepsilon_i}\right)=\frac{w_i}{A_i} \qquad (6)$$

In this equation, ε_i is the price elasticity of demand corresponding to i sort of Product.

As to demand, Harrod (1936) proposed the decreasing law about price elasticity of demand. In other words, the elasticity of demand will go down as income increases. This makes the demand on original products get saturated, but the demand on new products gradually go up, which leads to the diversification of demand on products. According to the descending tendency of Price Elasticity of Demand shown by converse tangent utility function, and the hypothesis about maximization of utility, it can be drawn:

$$\varepsilon_i=\varepsilon=\frac{n-1+(n+1)c^2}{2nc^2} \qquad (7)$$

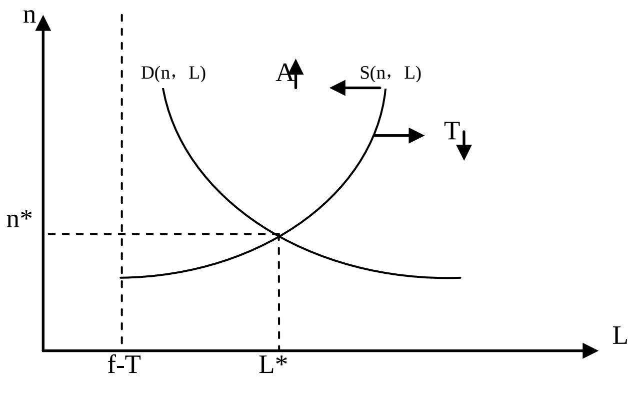
When equilibrium is reached, the degree of diversification presented by n* is decided by functions (5.8) and (5.9) together:

$$n=\frac{L-f+T}{f-(2L-f+T)\left[\frac{A(L-f+T)}{\bar{L}}\right]^2}+1 \qquad (8)$$

$$nL=(1-\alpha)\bar{L} \qquad (9)$$

In which, the total labor of two types of industry is fixed to be $\bar{L}$, (1-α) is the proportion of consumer's consumption in industrial products. According to the above equations, we can decide the degree of products diversification n* when other factors are in balance, as shown in Figure 5.

Figure 5. The determination of products diversification degree



The rise in A as a result of technological progress will move the curve S(n, L), which is determined by Equation (5.8) , towards the up-left, and the curve D(n, L) determined by Equation (5.9) will remain constant, then the degree of products diversification n* , determined by the intersection of these two curves, will increase. The reduction in transaction costs T will move the curve S(n, L) towards the bottom-right, and the curve D(n, L) will remain unchanged, thus the degree of products diversification n*, determined by the intersection of these two curves, will decrease.

Therefore, with the assumption that the consumer's demand is diverse, the progress in technology will cause the increase of product sales, the decrease in price elasticity of demand, and the rise in manufacturer's revenue, and therefore promotes the diversification of production. Moreover, the reduction in transaction costs will force the manufacturer to concentrate on the production of products with comparative advantages and outsource product with low productivity, then consequently promote product specialization.

This explanation also coincides with the view of "New Schumpeter" school in Evolutionary Economics about the co-evolution of technology, policy and industry structure. For example, in the early stage of information and communication technology revolution, the speed of technology development will be higher than that of transaction costs decrease, thus the regional industry structure turns to be more diverse; However, with time passing by, when the technology development slows down, the renovation in policy for full commercialization of technology will lead to the decrease in transaction costs. Its speed will gradually exceed the growth rate of technological progress, and therefore the industrial specialization will be achieved again.[9]

In summary, the evolution of regional industry specialization will be influenced by the demand, supply and policy of this region. First, as far as demand is concerned, as income level rises, the elasticity of demand for products will decrease. This will lead to gradual saturation in the demand of existing products and the increase in the demand of new products, which then cause the diversification of products in demand. Second, as far as supply is concerned, the advancement in technology companies the diversification of demand, improves the manufacturer's profits, and induced the entry of new manufacturers and new products, which then promotes the degree of production diversification. Third, in the aspects of policies, the reduction in transaction costs will force the manufacturer to concentrate on the production of products with comparative advantages and outsource product with low productivity, then consequently promote product specialization.

4. 2. Macro Model of the Evolution of Regional Industrial Specialization

The simplified Two-Sector Model explains the evolutionary mechanism of regional industrial specialization. But as a matter of fact, transaction cost can be divided into many kinds, besides, there are also many other critical variables that take effects on regional industrial specialization: Increasing return to scale, price level, wage rate, labor quantities and material capital stock. For this reason, we loosen the hypothesis of the micro model above, so as to concern the possible decisive factors of regional industrial specialization thoroughly.

In terms of macro impacting factors of regional industrial specialization, Dornbusch, Fischer and Samuelson (1977) found a Ricardian Trading Model which contains a continuous commodity set. Further, by studying the variation of scopes of borderline commodity produced by two countries, we explain the influence of technical progress on model of regional specialized production, and the mechanism that declining of transport cost and tariff will lead the non-trade commodity to trade commodity. When analyzing the effects of technical progress, the model doesn't bring in specific production function, only but to conduct a qualitative analysis on the effects of labor productivity's advancement. The analysis of transport cost and tariff's effects is only restricted to the extension of tradable commodities, but not the effects on regional industrial specialization in a detailed way.

In order to carry out a theoretical analysis on the impetus of Regional Industrial Specialization's evolution, we fit into the model above a Cobb-Douglas production function that is Hicks Neutral, and bring in the concept of "transaction costs" （neoinstitutional economics）to substitute the transport cost and tariff in the original model, thus found a statistical model of regional industrial structures' evolution in opening economy. The model makes an analysis that: During the development of economy, how technical progress and decrease of transaction cost affects the amount of regional industrial sectors, which is measured by the scope of borderline commodity production, and the variation of regional industrial specialization lead by the effects.

In order to coincide with the hypothesis of the model, we make an order of the continuous commodity set (defined in the interval of 0 to 1), according to the comparative advantage of the products manufactured in region *i* at home, compared to the products that produced abroad. Assuming product *z* is matched to each point in the interval, the input of labor is *a (z)* when producing a unit commodity z in region *i* at home. Since *a (z)* is monotonically increasing in the interval [0, 1], we assume that $a(z)=z^{\theta}$, the value of Constant $\theta>1$. Owing to the difference of relative labor productivity $\frac{1}{a(z)}$ between two regions, the scope of products, which is specialized produced in region i at home, is $z\in[0,\tilde{z}]$, the scope of the products that produced aboard is $z\in[\tilde{z},1]$. Among the production, the set, which is in the interval of $0<\tilde{z}<1$, is claimed as borderline commodity. We assume that each of the regional manufacture *i* at home apply Cobb-Douglas Production Function to produce Commodity *z*:

$Q(z)=\frac{A}{a(z)}K(z)^{\alpha}L(z)^{\beta}$, Q stands for Quantity, A stands for Total Factor Productivity in Region i, L and K are Labor and Capital Inputs, $0<\alpha<1$, $0<\beta<1$. The Profit Function of producing and selling Product z by the manufacture in Region i is:

$$\pi(z)=p(z)\frac{A}{a(z)}K(z)^{\alpha}L(z)^{\beta}-wL(z)-rK(z) \tag{10}$$

In this function, *p(z)* is the price of product *z* in region *i*, *w* and r are wage rate and rate of interest. When the profit is maximized:

$$\frac{\partial \pi}{\partial L} = p(z)\frac{A\beta}{a(z)}K(z)^{\alpha} L(z)^{\beta-1} - w = 0 \tag{11}$$

$$p(z) = \frac{z^{\theta} wL(z)^{1-\beta}}{A\beta K(z)^{\alpha}} \tag{12}$$

We assume that P* is the price of the product *z'* which is sold by manufacturers abroad, and the price is designated exogenously. During the course of shipping to region *i*, the total transaction cost（including transport cost ）is T, the price of the foreign product *z'* which is shipped to region *i* homeland is:

$$p(z') = TP^{*} \tag{13}$$

For the critical good $\tilde{z}$, there exists a formula $p(z) = p(z')$. Combining the formula (12) and (13), we can get the equation:

$$\tilde{z} = (\frac{\beta}{w} ATP^{*})^{\frac{1}{\theta}} K(z)^{\frac{\alpha}{\theta}} L(z)^{\frac{\beta-1}{\theta}} \tag{14}$$

Assuming each product is attached to a specific industry, the regional industrial specialization index (representing the diversification of products in one area) can be set as: $Div = \gamma \tilde{z}$, the constant $\gamma > 0$. So, we take the logarithm of the equation above as follows:

$$\ln Div = \ln\left(\gamma \tilde{z}\right) = \frac{1}{\theta}\ln A + \frac{1}{\theta}\ln T + \frac{1}{\theta}\ln P^{*} + \frac{\alpha}{\theta}\ln K + \frac{\beta-1}{\theta}\ln L - \frac{1}{\theta}\ln w + \frac{1}{\theta}\ln \beta + \ln \gamma \tag{15}$$

The regional industrial specialization index is in inverse proportion to the index of regional industrial diversification, that is Spec=1/Div, substituting the （15）by it, then we get the equation:

$$\ln Spec = -\frac{1}{\theta}\ln A - \frac{1}{\theta}\ln T - \frac{1}{\theta}\ln P^{*} - \frac{\alpha}{\theta}\ln K + \frac{1-\beta}{\theta}\ln L + \frac{1}{\theta}\ln w - \frac{1}{\theta}\ln \beta - \ln \gamma \tag{16}$$

As is known from formulas (15) and (16), the higher the TFP A is in one area, the more the autarky products are produced, the more diversified the industrial structure tends to be. While if the transaction cost T is the lowered, it tends to purchase products of non comparative advantaged products in trade, in that the regional industrial structure tends to be specialized. The lower the local price is, the higher the non-local products' price P* is, consequently, the region tends to manufacture all sorts of products by itself but not purchasing products outside, and the degree of diversification of regional structures will be advanced. The more abundant the regional capital K is, the higher the degree of regional industrial diversification is; the more sufficient the labor L is, and the higher the degree of regional industrial specification is. Thirdly, the increase of wage in one area will promote the development of regional industrial specialization.

4. 3. Empirical test of provincial panel data in China

On the base of mathematical model, we conduct a further empirical research by panel data of each province, so as to test the theory. In order to prove the evolutionary mechanism of regional industry specialization, we simplify the formula (16) without concerning differences of price, capital stocks and labor stocks between regions. But adhesiveness takes effects when labor transfers, so the regional industrial specialization will last for a period temporal dimension. The paper takes the first-degree lagged item of Regional Industrial Specialization as the explanatory variable and fit it into a model. Therefore, we assume the econometric model like this:

$$\ln Spec_{it} = \beta_1 \ln Spec_{it-1} + \beta_2 \ln A_{it} + \beta_3 \ln T_{it} + \mu_i + \nu_{it} \qquad (17)$$

Among them, *i* stands for province, *t* stands for year, μ_i is unobserved individual difference of each province, ν_{it} is stochastic error term. Since lnA and lnT stands for Total Factors Productivity and the increasing rate of transaction cost, so it can be easily seen from the formula that: When the increase of Total Factors Productivity exceeds the speed of declining of transaction cost, the degree of regional industrial specialization will keep dropping. Conversely, when declining of transaction cost exceeds the increase of Total Factors Productivity, the degree of regional industrial specialization will rise up again. The measurement and calculation of A_{it} and T_{it} are stated as follows:

Measuring and calculating the first-order lagged item of regional industrial specialization $Spec_{it-1}$: We select the regional industrial specialization index of each province in China as the value of the first-order lagged item. As the adjustment of industrial structure is a long-time process, and there exists adhesiveness when labor transfers, the level of regional industrial specialization displays continuity as time goes by. For the analysis above, we assume that $0 < \beta_1 < 1$.

The measurement and calculation of TFP (Total Factors Productivity) A_{it}: Agreeing with Cobb-Douglas set in statistics model, we apply $\ln Y_t = \ln A_t + \alpha \ln K_t + \beta \ln L_t$ to calculate Total Factor Productivity of each province from 1988 to 2007. The data is detailed as follows: The total output Y_t is represented by actual GDP volume of each province; total input of labor is represented by records of employment population of each province; total capital stock K_t is represented by actual capital stock. The data source of the first two variables are cited from statistical yearbook in each province; data on actual capital stock is quoted from material object capital of each province from 1978-2006, which is computed by Bai, Hsieh and Qian (2006).

Referring to the calculating method, we supplement the actual capital stock of 2007. In estimating the output elasticity of capital and labor, the estimation is divided into two periods 1988-1997 and 1998-2007 chronologically, and the scope of estimation are divided into three parts geographically: East , Middle, West. The increasing of TFP helps a enterprise to produce more sorts of products, and promotes the level of regional industrial diversification. We assume $\beta_2 < 0$.

The measurement and calculation of transaction cost T_{it} : So far , there are many disputes about the definition of transaction cost, which makes it difficult to measure the transaction cost (Zhao, 2005). Among vast of definitions, the one proposed by Yang (2003) is comparatively thorough and reasonable. Since it not only take into account the institutional factors, but also include technical factors like transport cost, trading infrastructures and so on. The transaction cost is divided into two kinds: exogenous transaction cost which is incurred directly and indirectly during trading process; endogenous transaction cost, which take place when the sufficient division of labor is held up by speculating behavior. The direct exogenous transaction cost is the cost incurred by transport; the indirect exogenous transaction cost refers to the expense on trading infrastructures like computers and credit cards used in communication and transactions. Exogenous transaction cost can be cut down by improvement of commuting and informational infrastructures, while endogenous transaction cost can be reduced by arrangement of institution and contracts. Yang (1988) is the first to propose the definition of trading efficiency: When someone buys one-unit product, and he or she only gets *k* unit product, hence 1-k unit can be regarded as trading efficiency. By applying the converse correlation of transaction cost with transaction efficiency, Yang solves the problem when measuring and calculating transaction cost. In the following, we will analyze the factors that promote the transaction efficiency, and then measure the transaction cost.

Specifically speaking, we use the efficiency of transport infrastructures in designate provinces to measure the declining of exogenous transaction cost. We will apply the method used by Demurger (2001), that is aggregating the milometer of railway, highway and cruise way, and then the total number is divided by the size of territory of each province. Improvement of informational infrastructures also can promote the transactional efficiency greatly, and reduce the indirect exogenous transaction cost (Hendriks, 1999), and we refer to Liu and Hu (2010) for their method to measure it. In this method, they choose a comprehensive index, that is business volume of postal and telecommunication services, to reflect the efficiency of the informational infrastructures in each province. Hall and Jone (1999) point out that social infrastructure, which contains institution and governmental policy, promote the transactional efficiency greatly, and reduce the endogenous transaction cost. For China, Reform and opening-up are the most important innovation on institution, which reduce the endogenous transaction cost dramatically and promote the development of economy to a great extent. We choose the proportion of the employment of non state-owned enterprises to the total employment in this area to measure the degree of reform, and choose the proportion of total export-import volume and foreign direct investment (FDI) that contributes to GDP in this region, to reflects the degree of openness of this region, both of which are combined to measure the reduction of endogenous transaction cost. Finally, the 5 indicators above will be converted into scalars and their reciprocals will be calculated, then the 5 reciprocals will be averaged by the same weightiness to work out a comprehensive index of transaction cost.

According to the estimation function (17), we set hypothesis on coefficient as follows: β_1 represents the coefficient of the industrial specialization of previous period' s affection on that of next period ($0 < \beta_1 < 1$), β_2 represents coefficient of technology progress's influence on regional industry specialization ($\beta_2 < 0$), β_3 represents the coefficient of transaction cost reduction' effects on regional industry specialization ($\beta_3 < 0$). As the formula (17) will use the first-order lagged item of explained variable as one of the explanatory variables, it will makes the explanatory variable to be correlated with random error term. If standard random effects or fixed effects model is applied, the results would be biased. For the above reasons, this paper applies difference-GMM and system-GMM to estimate dynamic panel. By using difference method to model and setting Instrumental Variables of the difference value of the explanatory variables, we can work out the estimators of difference-GMM (Arellano and Bond, 1991). The difference-GMM estimation can effectively settle the problems like endophytism of explanatory variables and heteroscedasticity of residual. But the drawbacks of difference-GMM lie in that: it will incur informational loss of a few samples; and when the continuity of explanation keeps long, the effectiveness of instrumental variables will be weakened, as a result, the gradual effectiveness of the estimation will be affected. Arellano and Bover (1995) , as well as Blundell and Bond (1998) suggest choosing system-GMM, that is fitting the original function (containing variable level) and the function of one-ordered difference into one system. System-GMM can make use of the information in difference function and horizontal structure function simultaneously, and more often than not, is more effective than difference-GMM. When using GMM estimation, we still have to exam the effectiveness of instrumental variables and whether the residual item is Serial correlation (Bond, 2002). Estimation in the paper chooses Sargan Statistics to test effectiveness of instrumental variables. If it accepts the original hypothesis, the selection of instrumental variables are effective; Besides, we employ the first-order and second-order Serial correlation test AR(1) and AR (2) of first-order difference converted function to judge the residual items are Serial correlation. In terms of weightiness of the matrix, GMM can be further divided into one-step GMM estimation and two-step GMM estimation. Compared to one-step GMM estimation, the two-step GMM estimation is not probably disturbed by Garch. Therefore, the paper uses the two-step GMM to estimate the formula (17).

The results of estimation are shown in Table 5: The first four columns are the results of difference-GMM estimation, the later four columns are the results of system-GMM estimation. If Sargan Statistics are not significant, it can be inferred that the selection of instrumental variables are effective; if AR (1) test rejects the null hypothesis and AR (2) test accepts the null hypothesis , it can be inferred that the residual of the original function is serially uncorrelated.. The specific empirical results of each coefficient in the model are presented as follows:

The first-degree lagged item of Regional Industrial Specialization Index is significantly positive, which testify the hypothesis that adhesiveness takes effects when labor transfers, and the regional industrial specialization will last for a period temporal dimension.

The empirical results coincide with the expectation: The technology progress coefficients of East, Middle, West part of China, as well as the level of provinces are all significantly negative. It shows that the increasing of Total Factor Productivity promote the diversification of regional industrial structure in Eastern, Middle, Western Part of China.

Coinciding with the prediction, transaction cost coefficients of the Middle, Western China, as well as the provincial level, are all significantly negative. It shows that declining of transaction cost greatly promote the regional industry specialization in Eastern, the Middle, and Western part of China.

Table 5. The empirical results of factors that affect regional industry specialization

variable	Logarithm of regional industrial specialization index							
Estimation	Difference-GMM				System-GMM			
Region level	**Provincial**	Eastern	Middle	Western	**Provincial**	Eastern	Middle	Western
L.lnSpec	0.473^{***} (6.01)	0.331^{***} （4.12）	0.272^{***} （3.35）	0.742^{***} （9.72）	0.626^{***} (12.37)	0.293^{***} （3.13）	0.335^{***} （4.31）	0.697^{***} （10.35）
lnA	-0.137^{***} (4.65)	-0.023^{**} (2.40)	-0.251^{*} (1.81)	-0.177^{*} (1.69)	-0.209^{***} (3.85)	-0.086^{***} (2.96)	-0.170^{**} (2.47)	-0.192^{**} (2.51)
lnT	-0.212^{***} (6.80)	-0.175^{***} (5.42)	-0.177^{**} (2.23)	-0.281^{**} (2.43)	-0.247^{***} (7.11)	-0.132^{***} (4.02)	-0.146^{*} (1.77)	-0.237^{**} (2.38)
AR(1)	0.013	0.021	0.016	0.035	0.009	0.015	0.011	0.027
AR(2)	0.437	0.325	0.531	0.230	0.397	0.366	0.318	0.291
Sargan test	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Notice：Regression results are computed by STATA11.0. The results in the brackets are absolute value of T statistic, *, **, *** indicates the results have passed the test of significance at the significant level of 10%, 5%, 1%. AR(1), AR(2) and Sargan test's reporting value is the one that meets probability value>z

5. The causes analysis of the unbalanced regional industry specialization development of China

In the last section, we only analyzed the interaction of technological progress and transaction cost which causes the U-shaped curve evolution law of regional industry specialization. But which detailed factors influence the regional industry specialization development and what the extent they influence? We use econometrics method to measure the influence of relevant factors, and sort their orders by their contribution to the unbalanced regional industry specialization development of China, in order to make policies to promote the regional industry specialization development scientifically.

As described in part III of this paper, in the rising phase of the U-shaped curve, the gap between the provincial level of regional industry specialization is widening continuously. Based on the macro-econometric model, we used panel data of 12 Chinese provinces in 1999-2007 to analyze the detailed factors which cause the unbalanced regional industry specialization development of China by regression-based inequality decomposition method empirically and systematically. We still use formula (16) to set the econometric model:

$$\ln Spec_{it} = \beta_1 \ln Spec_{it-1} + \beta_2 \ln A_{it} + \beta_3 \ln T_{it} + \beta_4 \ln P_{it}^{*} + \beta_5 \ln K_{it} + \beta_6 \ln L_{it} + \beta_7 \ln w_{it} + \mu_i + \nu_{it} \quad (16)$$

In this function, i represent different provinces, t representative different years, μ_i is the unobservable individual differences in the cross-section unit, ν_{it} is stochastic error term. The measurement and calculation of the explanatory variables are stated as follows:

1. Measuring and calculating the first-order lagged item of regional industry specialization $Spec_{it-1}$: We select the regional industry specialization index of 12 provinces in China from 1998 to 2006 as the value of the first-order lagged item. As the adjustment of industrial structure is a long-time process, and there exists adhesiveness when labor transfers, the level of regional industry specialization displays continuity as time goes by. For the analysis above, we assume that $0 < \beta_1 < 1$.

2. The measurement and calculation of TFP (Total Factors Productivity) A_{it}, total capital stock K_{it} and total input of labor L_{it}: Agreeing with Cobb-Douglas set in Statistics Model, we apply $\ln Y_t = \ln A_t + \alpha \ln K_t + \beta \ln L_t$ to calculate Total Factor Productivity of each province from 1999-2007. The data is detailed as follows: The total output Y_t is represented by actual GDP of each province, total input of labor L_t is represented by records of employment population of each province, total capital stock K_t is represented by actual capital stock. The data source of the first two variables are cited from statistical yearbook in each province, data on actual capital stock is quoted from material object capital of each province from 1978-2006, which is computed by Bai, Hsieh and Qian (2006). Referring to the calculating method, we supplement the actual capital stock of 2007. The increasing of TFP helps a enterprise to produce more sorts of products, and promotes the level of regional industrial diversification. We assume $\beta_2 < 0$. The increasing of capital stock helps a enterprise to produce other sorts of products, and promotes the level of regional industrial diversification. We assume $\beta_5 < 0$. The increasing of main industry output helps workers to assemble in this area, and promotes the level of regional industry specialization. We assume $\beta_6 < 0$.

3. The measurement and calculation of transaction cost T_{it}: So far , there are many disputes

about the definition of transaction cost, which makes it difficult to measure the transaction cost (Zhao, 2005). Among vast of definitions, the one proposed by Yang (2003) is comparatively thorough and reasonable. Since it not only take into account the institutional factors, but also include technical factors like transport cost, trading infrastructures and so on. The transaction cost is divided into two kinds: exogenous transaction cost which is incurred directly and indirectly during trading process; endogenous transaction cost, which take place when the sufficient division of labor is held up by speculating behavior. The direct exogenous transaction cost is the cost incurred by transport; the indirect exogenous transaction cost refers to the expense on trading infrastructures like computers and credit cards used in communication and transactions. Exogenous transaction cost can be cut down by improvement of commuting and informational infrastructures, while endogenous transaction cost can be reduced by arrangement of institution and contracts. Yang (1988) is the first to propose the definition of trading efficiency: When someone buys one-unit product, and he or she only gets *k* unit product, hence 1-k unit can be regarded as trading efficiency. By applying the converse correlation of transaction cost with transaction efficiency, Yang solves the problem when measuring and calculating transaction cost. In the following, we will analyze the factors that promote the transaction efficiency, and then measure the transaction cost.

Specifically speaking, we use the efficiency of transport infrastructures in designate provinces to measure the declining of exogenous transaction cost. We will apply the method used by Demurger (2001), that is aggregating the milometer of railway, highway and cruise way, and then the total number is divided by the size of territory of each province. Improvement of informational infrastructures also can promote the transactional efficiency greatly, and reduce the indirect exogenous transaction cost (Hendriks, 1999), and we refer to Liu and Hu (2010) for their method to measure it. In this method, they choose a comprehensive index, that is business volume of postal and telecommunication services, to reflect the efficiency of the informational infrastructures in each province. Hall and Jone (1999) point out that social infrastructure, which contains institution and governmental policy, promote the transactional efficiency greatly, and reduce the endogenous transaction cost. For China, Reform and opening-up are the most important innovation on institution, which reduce the endogenous transaction cost dramatically and promote the development of economy to a great extent. We choose the proportion of the employment of non state-owned enterprises to the total employment in this area to measure the degree of reform, and choose the proportion of total export-import volume and foreign direct investment (FDI) that contributes to GDP in this region, to reflects the degree of openness of this region, both of which are combined to measure the reduction of endogenous transaction cost.

4. The measurement and calculation of price in other places P_{it}^*: We use the ΔCPI which is gap between national consumer price index (CPI_t) and provincial consumer price index (CPI_{it}) to reflect the difference between the price in other places and the local price. Along with the rising ΔCPI , the goods which are made in other places will be sold more expensive than the goods which are made locally. The local manufacturers will tend to expand the product range to earn

excess profits. Therefore, the level of regional industry specialization is negatively correlated with the price in other places, we assume $\beta_4 < 0$.

5. The measurement and calculation of wage w_{it}: We use the total wages divided by the provincial employment to measure the level of regional average wage $Wage$. According to the two regional models Krugman (1991) described: If the number of local manufacturers is more, the workers and consumers will be able to buy the necessary goods with lower transportation costs, so the local price index will be lower and the local real wage will be higher. Thus the workers will prefer to live and work in this region, the industry will further concentrate in this region. Therefore, the level of regional industry specialization is positively correlated with the local wage, we assume $\beta_7 < 0$.

Our econometric model (16) can be further described in detail as :

$$\begin{aligned}\ln Spec_{it} &= \beta_1 \ln Spec_{it-1} + \beta_2 \ln TFP_{it} + \beta_{31} \ln Transport_{it} + \beta_{32} \ln Information_{it} \\ &+ \beta_{33} \ln Reform_{it} + \beta_{34} \ln Trade_{it} + \beta_{35} \ln FDI_{it} + \beta_4 \ln \Delta CPI_{it} + \beta_5 \ln K_{it} \\ &+ \beta_6 \ln L_{it} + \beta_7 \ln Wage_{it} + \mu_i + \nu_{it}\end{aligned} \quad (18)$$

As the formula (18) will use the first-order lagged item of explained variable as one of the explanatory variables, it will makes the explanatory variable to be correlated with stochastic error term. We still apply difference-GMM and system-GMM to estimate dynamic panel. In terms of weightiness of the matrix, GMM can be further divided into one-step GMM estimation and two-step GMM estimation. we still apply the two-step GMM to estimate the formula (18).

The results of estimation are shown in Table 6: The first four columns are the results of difference-GMM estimation, the later four columns are the results of system-GMM estimation. If Sargan statistics are not significant, it can be inferred that the selection of instrumental variables are effective, if AR (1) test rejects the null hypothesis and AR (2) test accepts the null hypothesis , it can be inferred that the residual of the original function is serially uncorrelated.

Based on this regression results, we use the regression-based inequality decomposition method Wan (2002) proposed, which combined the regression equation with the Shapley value decomposition principle together, to sort order of these factors. Regression equation relates the target variable such as income, consumption, with its determinants, such as human capital, family characteristics, professional nature, sex, age, etc. to establish a quantitative relationship. Shapley value theory is based on cooperative game theory, with which the inequality of the target variable can be decomposed into the contribution of the determinants. Our inequality decomposition results are based on the system GMM estimates.[10]

Table 6. The empirical results of factors that affect regionalindustry specialization at level of 12 provinces

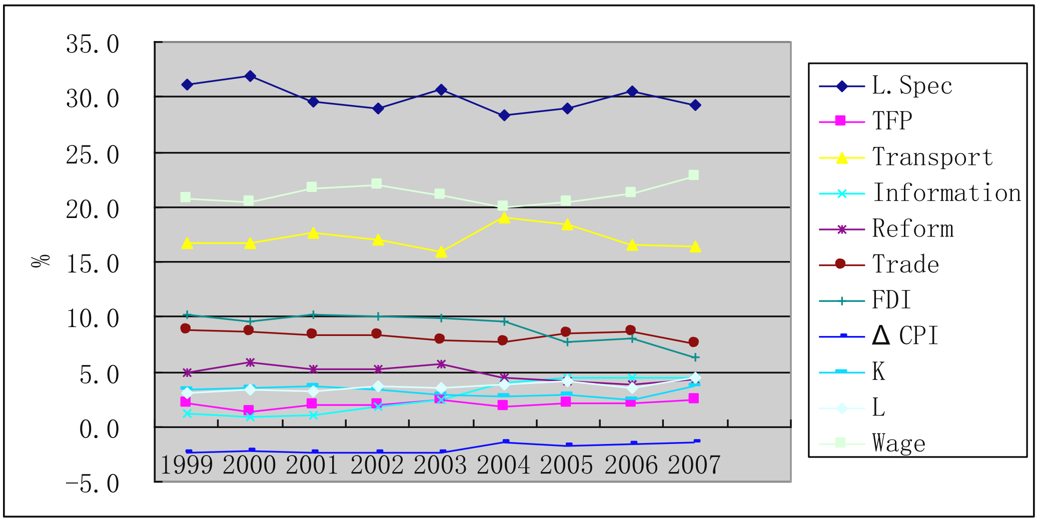
Dependent variable	Logarithm of regional industry specialization index

Estimation method	Difference-GMM		System-GMM	
Independent variable	coefficient	z-value	coefficient	z-value
L.lnSpec	0.275***	6.71	0.309***	7.59
lnTFP	-0.107**	-2.29	-0.085*	-1.80
lnTransport	0.013**	2.48	0.034***	2.86
lnInformation	0.021*	1.85	0.047**	2.35
lnReform	0.014*	1.79	0.025*	1.83
lnTrade	0.116***	5.56	0.076***	4.93
lnFDI	0.042***	3.08	0.031***	3.88
lnΔCPI	-0.019***	-2.79	-0.023***	-3.04
LnK	0.142***	4.12	0.160***	4.71
lnL	0.091***	3.31	0.077***	3.43
lnWage	0.026***	3.19	0.038***	5.61
Abond test for AR(1)	0.027		0.016	
Abond test for AR(2)	0.505		0.482	
Sargan test	1.00		1.00	

Notice：Regression results are computed by STATA11.0. *, **, *** indicates the results have passed the test of significance at the significant level of 10%, 5%, 1%. AR(1), AR(2) and Sargan test's reporting value is the one that meets probability value>z

As Table 6 shows, with the exception of the capital stock coefficient is positive which is not accord with the theoretical model, the others are consistent with the theoretical model. This may be caused by Chinese banks prefer lending to state-owned enterprises and the reform object of Chinese state-owned enterprises requires them to peel off secondary business in recent years. With the specialization level of state-owned enterprises increasing continuously, the level of regional industry specialization is positively correlated with the capital stock. According to the estimated regression coefficients, we use the software Wan Guanghua provided to carry out the inequality decomposition. As shown in Figure 6, decomposition result visually describes the inequality contribution percentage of different economic factors. According to the contribution of regional industry specialization inequality, the original level of industrial specialization, wage, transport infrastructure, foreign trade and foreign direct investment are the most important factors which lead to inequality of Chinese regional industry specialization in order.

Figure 6. The contribution of all factor on inequality of regional industry specialization



From the above decomposition, we can draw the following policy conclusions:

First, inequality contribution of the original level gap of regional industry specialization is about 30%, which reveals the adjustment of the regional industrial structure will be a long process. Provincial government should promote the industrial structure adjustment based on existing industries, and reduce regional industry specialization inequality.

Second, the provincial difference in wage level is the second factor, it contributes about 20% of the regional industry specialization inequality. Income gap between provinces has became an important factor which caused regional industry specialization inequality. In this regard, the government should improve the income distribution mechanism, strengthen transfer payments to the backward areas, improve the income of local workers and reduce the inequality between regional industry specialization.

Third, transportation infrastructure condition, followed by wage, contributes about 15% of the regional industry specialization inequality. This shows that the decreasing transaction cost which is reduced by the transport infrastructure improvement will narrow regional industry specialization gap. Thus, the government should increase transport infrastructure investment in backward areas.

Fourth, foreign trade and foreign direct investment also are important factors which cause regional industry specialization inequality. The former contributes about 7%-9%, the latter contributes about 6%-10%. Thus increasing the foreign trade of inland province and attracting more foreign investment also could promote the industrial structure adjustment and regional industry specialization.

Fifth, the progress of reform, the amount of labor, information infrastructure, capital stock, total factor productivity and price level impact regional industry specialization inequality less than the others, their contributions are less than 5%. This shows that the government further advance reform, increase labor mobility, improve information infrastructure, increase the capital stock and improve total factor productivity can also promote the balanced development of regional industry specialization, but the effect is less.

6. Conclusions and policy suggestions

In this article, we tried to systematically reveal the general evolutionary laws of regional industrial specialization in the process of economic development from theoretical and empirical view. And we concluded the deep economic reasons behind the evolution of Chinese regional industrial structure. We found that in the process of economic development, the evolution of Chinese regional industry specialization is in accord with U-shaped curve, that is the industrial structure firstly diversified in the early stages of economic development, then specialized in the latter part of economic development. Whether it is in the national and provincial level, or in the eastern, central and western regions level, this law is set up, and the stage of development of regional industry specialization is closely related with the level of GDP per capita.

Why the evolution law of regional industry specialization will be U-shaped curve? Our explanation is that technological progress promotes regional industrial diversification, while reducing transaction cost promotes regional industry specialization in the process of economic development. When the growth rate of technological progress is faster than the decline rate of transaction cost, the regional industrial structure will be diversified. Otherwise, the regional industry specialization will become the mainstream of economic development. We give a simple mathematical model of the micro evidence on this conclusion, and give a detailed mathematical model to explore the various actual determinants of unbalanced Chinese provincial industrial specialization development. Based on the decomposition of inequality, we found that the original level of industrial specialization, wages, transport infrastructure, foreign trade and foreign direct investment are the most important determinant of differences in regional industry specialization .

According to the results, we can get three policy implications: First, for the specialization has became the inevitable trend of a higher stage of economic development, the policy makers in the eastern, central and western regions must pay attention to the specialized industrial structure adjustment direction. In particular, central and western regions, which have just entered the increased regional industry specialization period, must adjust their development strategy to catch up with the eastern and coastal provinces and promote regional industrial structure adjustment and optimization. Second, along with China's income distribution has attracted increasing attention of the central government, if we can reduce the regional wage differences and promote regional industrial structure specialization and optimization at the same time, it would be killing two birds with one stone policy prescriptions. Third, eastern provinces have a first-mover advantage in regional industry specialization with the better transportation infrastructure, more foreign trade and more FDI. And the original level of regional industry specialization has decisive influence on the regional industry specialization development, which further strengthen that advantage. So the central government must give a larger amount of infrastructure investment, more preferential import, export and investment policies to the inland provinces, in order to help them get rid of the "low-level equilibrium trap" in the regional industry specialization development, and promote regional industrial structure adjustment and optimization effectively.

References

[1] Antonelli C., The Economics of Innovation, New Technologies and Structural Change, Taylor and Francis Group, 2003.

[2] Arellano M. and Bond S., Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations [J], *Review of Economic Studies*, 1991, 58(2):277-297.

[3] Arellano M. and Bover O., Another Look at the Instrumental Variable Estimation of Error-Components Models [J], *Journal of Econometrics*, 1995, 68(1): 29-51.

[4] Bai C.E., Hsieh C.T. and Qian Y, The Return to Capital in China [J], *Brookings Papers on Economic Activity*, 2006, 2: 61-88.

[5] Blundell R. and Bond S., Initial Conditions and Moment restrictions in Dynamic Panel Data Models [J], *Journal of Econometrics* ,1998, 87(1):115-143.

[6] Bo W., Externalities and Industrial Economic Growth——Evidences from Chinese Provincial Panel Data [J], *China Industrial Economy*, 2007, 1:37-44.

[7] Bond S., Dynamic Panel Data Models: A Guide to Micro Data Methods And Practice[J], *Portuguese Economic Journal*, 2002, 1(2): 141-162.

[8] Cao J., Technological progress impacts on product diversification [J], economic science, 2008, 1:104-113.

[9] Cleveland W. S., Robust Locally Weighted Regression and Smoothing Scatterplots [J], *Journal of the American Statistical Association*, 1979, 74(368):829-836.

[10] Démurger S., Infrastructure Development and Economic Growth: An Explanation for Regional Disparities in China? [J] *Journal of Comparative Economics*, 2001, 29(1):95-117.

[11] Dornbusch R., Fischer S. and Samuelson P.A., Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods[J], *American Economic Review*, 1977, 67(5):823-839.

[12] Fan F., The Measurement of Regional Specialization [J], *Economic Research*, 2007, 9:71-83.

[13] Fan J., Market Integration, Regional Specialization and Tendency of Industrial Agglomeration: An Implication for Regional Disparity [J], *Social Sciences in China*, 2004, 6:39-51.

[14] Feldman M.P. and Audretsch D.B., Innovation in cities: Science-based diversity, specialization and localized competition [J], *European Economic Review*, 1999,43(2):409-429.

[15] Fujita M., Krugman P. and Venables A. J., *The Spatial Economy: Cities, Regions, and International Trade*, [M] Cambridge: The MIT Press, 1999, 239-262.

[16] Glaeser E.L., Kallal H.D., Scheinkman J.A., and Shleifer A. Growth in Cities [J], *Journal of Political Economy*, 1992, 100(6):1126-1152.

[17] Hall R.E. and Jones C.I., Why Do Some Countries Produce So Much More Output Per Worker Than Others?[J] *Quarterly Journal of Economics*, 1999,114(1):83-116.

[18] He C. and Xie X., Geographical Concentration and Provincial Specialization of Chinese Manufacturing Industries [J], *Geographical Journal*, 2006 , 2:212-222.

[19] Hendriks P., Why Share Knowledge? The influence of ICT on Motivation for Knowledge Sharing [J], *Knowledge and Process Management*, 1999, 6(2):91-100.

[20] Huang X. and Shu Y., Openness, Industrial Specialization and China's Economic Growth [J], *International Trade Issue*, 2007, 12:11-17.

[21] Imbs J. and Wacziarg R., Stages of Diversification [J], *American Economic Review*, 2003, 93(1):63-86.

[22] Kim S., Expansion of Markets and the Geographic Distribution of Economic Activities: The Trends in US Regional Manufacturing Structure, 1860-1987 [J], *Quarterly Journal of Economics*, 1995, 110(4):881-908.

[23] Li J., Ran G. and Wan G. Study on the Disparities of Regional Financial Development in China: Based on the Theory of Division of Labor and Decomposition Method of Shapley Value [J], *Economic Research*, 2007, 5:42-54.

[24] Lin X., The empirical research of Chinese provincial industry specialization: 1988-2002 [J], *Shanghai Economic Research*, 2007 ,1:3-28.

[25] Malizia E. E.and Ke S., The Influence of Economic Diversity on Unemployment and Stability [J], *Journal of Regional Science*, 1993, 33(2):221-235.

[26] Sen A. and Foster J.E., *On Economic Inequality*, [M] Oxford: Oxford University Press, 1997, p.31.

[27] Shi P. and Zheng J., Innovation and Industry Specilization Transition Dynamics: The Case of Jiangsu [J], *Industrial Economics Research*, 2007, 1:43-48.

[28] Wan, G., Regression-based Inequality Decomposition: Pitfalls and A Solution Procedure, WIDER Discussion Paper, 2002.

[29] Wan G., Inequality Measurement and Decomposition: A Survey [J], *China Economic Quarterly*, 2008, 8(1):347-368.

[30] Wan G., Fan B. and Lu M. Analysis of Inequality in Chinese Innovation Capability—— Regression-based Decomposition Method [J], *World Economy*, 2010, 2:3-14.

[31] Yang X., A Microeconomic Approach to Modeling the Division of Labor Based on Increasing Return to Specialization [D], Ph.D. Dissertation, Department of Economics, Princeton University, 1988.

---

[1] See for instance Sen and Foster (1997).

[2] According to Cleveland's recommendations, in the following fitting, we select f = 0.5, that is we use 50% of the total sample which is near the central point in each sub-sample regression.

[3] From 1988-1992, the data is the industrial net output indicators. The data for 1995, 1996 and 1998 is not available. data for 2004 is from "China Economic Census Yearbook 2004" . Criteria for the classification of China's industrial categories in 1988, 1993 and 2003 occurred in large changes. In order to ensure the continuity of indicators and statistical consistency, we retain only the 1988-2007 data in continuous sectors. On this basis, we combined "oil-processing industry" with "coking, gas and coal products industry" into "oil processing and coking industry", and combined "ordinary machinery manufacturing" with "special equipment manufacturing" into the "machinery manufacturing industry ".Then we remove " electricity and heat production and supply industry ",which is a non-tradable sectors, remain a total of 20 industries data.

[4] The actual variables whose unit is RMB is measured by constant prices (1985=100). The nominal variables whose unit is RMB is measured by current prices.

5 The lowest point abscissa of quadratic curve Gini = α + β1(GDP) 2 + β2GDP is GDP =- β2/(2β1).

[6] In fact, this finding consists with the phenomenon which the share of industrial output changes earlier than the share of employment in Chinese industrial structure adjustment .

[7] Penn World Table version 6.3 see http://pwt.econ.upenn.edu.

[8]Eastern provinces: Beijing, Tianjin, Hebei, Liaoning, Shanghai, Jiangsu, Zhejiang, Fujian, Shandong, Guangdong and Hainan; central provinces: Shanxi, Jilin, Heilongjiang, Anhui, Jiangxi, Henan, Hubei, Hunan; western provinces: Neimenggu, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Tibet, Shaanxi,

Gansu, Qinghai, Ningxia and Xinjiang. In addition, a large number of data is missing in the Tibet and Hainan, so we discard it. Then we put Chongqing data into Sichuan. In the sample period, the GDP of Neimenggu is significantly higher than other western provinces, so we discard the data when study the western provinces. So this paper only covers the sample data in China's 28 provinces.

[9] See for instance Antonelli (2003).

[10] We thank Professor Wan Guanghua for providing regression-based inequality decomposition software. The version is WIDERdecom20041126.