CARPATHIAN JOURNAL OF FOOD SCIENCE AND TECHNOLOGY

journal homepage: http://chimie-biologie.ubm.ro/carpathian_journal/index.html

EMPIRICAL STUDY ON CHINA DAIRY INDUSTRIAL CLUSTER AND INFLUENCE FACTORS-BASED ON PROVINCIAL PANEL DATA SPATIAL ECONOMETRIC ANALYSIS

Zhiming Zhang[1*]

[1]*Capital University of Economics and Business, Beijing, China*
*Corresponding author: *cgfwzx@126.com*

Article history:
Received:
22 February 2016
Accepted in revised form:
29 May 2016

Keywords:
Dairy industry;
Industry cluster;
Spatial econometric model;
Spatial dependency;
Influence factors

ABSTRACT

The paper explores and analyzes China's dairy industrial cluster and influence factors by the employments of spatial lag model and a spatial error model related to the fixed effect and random effect and with spatial panel data of 29 provinces from the year of 2009 to the year of 2013 as samples. In conclusion, China's provincial dairy industry has strong spatial dependency and a positive spatial correlation; the degree of industrial cluster has a positive correlation with resource endowment, external economic conditions, industrial profitability as well as governmental supports, but has a negative correlation with economic foundation; and the labor cost does not play an obvious role in industrial cluster.

1. Introduction

With the development of dairy industry, many dairy enterprises have overcome the geo-spatial limit. It is gradually obvious for the dairy spatial cluster. Dairy enterprises quantity has increased to 658 from only a few at the beginning. And these dairy enterprises' geographical occurrence is mainly in north of China, and northwest of China. According to the statistics from National Bureau of Statistics, the dairy products output of the year 2013 grew by 5.15% over the same period of 2012 to 26,980,300 tons. And the liquid milk output grew by 7.01% over the same period of 2012 to 23,559,700 tons. The turnover of the dairy enterprises with scales rose by 14.16% over the same period of 2012 to RMB 283.159 billion. With the fast development of dairy industry, how is the correlation of China dairy industrial cluster spatial considering the industrial cluster spatial spillovers? What are the influence factors of cluster? By using spatial econometric analysis, the paper is try to construct China dairy industrial spatial econometric model to do empirical analysis on the influence factors, and to interpret the spatial spillover and heterogeneity of cluster, so as to provide theoretical reference for the government to make reasonable dairy industrial development policies. The research on the dairy industrial cluster and influence factors are in its beginning stage both at home and abroad and most of the researches are on the dairy industrial cluster measurement. Foreign scholars did more research on industrial cluster. As for the basic theory researches, new classic trading school believe the endowment of natural resources, labor, technology and others influenced the industrial cluster(Ohlin B, 1968). New economic geography school put the external scale economy into the industrial cluster analysis, and thinks a industrial cluster will be promoted due to the existence of external scale economy (Henderson JV, 1974; Fujita M.A, 1988). Some of the foreign

scholars did researches on a series of economic problems by using the spatial econometric model related theories. For instance, Lesage (1999) found that China provincial economy growth exists obvious spatial cluster ways by using spatial econometric model research. Meanwhile, the economy growth will influence the economy growth of other countries. It proved the spillovers of economy growth (Easterly, 1988). Tschoegl (2000) believes that external scales economy will promote the finance institutions to choose the certain areas. This area will be more advantageous when many finance institutions clustered here. Dairy industry cluster and influence factors are researched in foreign countries less than China.

Some of Chinese scholars did a series of researches on dairy industrial cluster. On the one hand, it is the research on the provincial dairy industrial cluster in all over China. China dairy clustered situation and spatial layout are the two important aspects which are influencing economic achievement and industrial competition. China dairy industry cluster degree and spatial cluster degree are not high (Hua, 2007). China dairy industry is clustered mainly in northeast China, Inner Mongolia, and north China. At the same time industry cluster showed marked enhancement in the dairy industry growth (Cheng, 2012). On the other hand, it is the research on some provinces dairy industrial cluster. Inner Mongolia dairy industry, livestock producing ability, returns to scale, industry investment location choice are the decisive factors of industry cluster. Livestock producing ability, industry cluster degree, the scale of original industry fixed assets are the decisive factors of industry investment location choice (Li, 2008). It also analyzed the development trends of Inner Mongolia animal products industry cluster (Li and An, 2008). Chinese scholars introduced spatial econometric model into the research on industry cluster, for example, on textiles industry, finance industry and others (Niu and Jiang, 2011; Ren et al., 2010).

The above researching results promoted the development of dairy industry cluster theory. However, the researches on China dairy industry cluster didn't take the correlation between the industry cluster and spatial geography location into consideration, neglected the space dimension heterogeneity, and used few econometric models in the empirical research on dairy industry cluster. There is significant difference among the different provinces. The traditional regression model can't effectively estimate the spillovers of the dairy industry cluster. Therefore, the paper put the spatial econometric model into the research on China 29 provinces dairy industry cluster influence factors, based on the provincial panel data, to explore the deep cause of dairy industry cluster.

2. Materials and methods

Theory hypothesis

Hypothesis 1: China dairy industry has spatial dependency and spatial spillovers.

Industry cluster refers to the same industry is highly concentrated in a certain geography location, and the industry capital clusters gradually in the spatial range. Since 19th century, the economists Alfred Marshall, Weber, Porter, Hoover have done deep analysis on industry cluster. Chinese scholars Zhang Xueliang, Zhao Liangshi, Sun Qinggang, Liu Hedong, Niu Honglei, Ren Yinghua respectively found that China infrastructure, utilization ratio of water resource, energy intensity, regional innovation, textile industry cluster, finance industry cluster all have the spatial dependency and spatial spillovers. There are few studies on China dairy industry cluster. The large scaled dairy enterprises like Mengniu, Yili, Bright, Firmus, Longdan, Wandersun have spatial geography proximity. Therefore, the paper makes the hypothesis that China dairy industry cluster has spatial dependency and spatial spillovers.

Hypothesis 2: Resources endowment namely the availability of raw materials is the

core factor of influencing dairy industry cluster, and is promoting effectively the dairy industry cluster.

New classic trading theory holds the opinion that the natural resources endowment advantage decides the industry location cluster [1]. For the industry that relies heavily on the raw materials, resources endowment is the key factor of industry location selecting, and is the foundation of the industry cluster process. Dairy industry needs raw milk as the raw materials. Raw milk is with short fresh-keeping time, and is perishable. Dairy industry would treat the raw milk availability as the first important factor to choose the industry location. Therefore, the paper put forward the hypothesis that the availability of raw materials is the key factor of influencing dairy industry cluster.

Hypothesis 3: Labor cost will significantly influence the dairy industry cluster. There is negative correlation between labor cost and dairy industry cluster.

New classic trading theory believes that except the natural resources endowment, labor force, technology, and other external resources endowment are the important factors of influencing industry cluster. Wang (2010) found that many industries clustered in coastal region caused a lot of middle east China countryside labor force transferring to the east China. Labor transfer will further promote the industry cluster. As the "soft production factor", labor force is the important factor of improving manufacturing industry competition. The cheap labor force is the cost advantage of local industry development, and promotes industry cluster. Dairy industry belongs to manufacturing industry. The availability of labor force and labor force cost will influence the dairy industry cluster. Generally, the less of the labor cost, the more attractive it will be for the manufacturing industry cluster.

Hypothesis 4: There is positive correlation between industry scale levels and dairy industry cluster.

The external economic theory, which was put forward by economist Marshall in 1890, and developed and perfected by scholar Krugman, believes that larger scaled industry is more efficient in manufacturing than the smaller scaled industry. The expansion of industry scale will cause the increase of the industry profit, thus the same industry and its supporting department will cluster in one or several places. Represented by Henderson [2] and Fujita [3], the new economic geography school thinks that the average cost decreased due to the increasing return to scale, and it increased the competition in further expanding scales, and it promoted the industry highly clustered. The current scale level of dairy industry will decide whether it will form the external scale economy. The larger scale level, the more helpful it will be for the dairy industry clustering in this place.

Hypothesis 5: There are highly positive correlation between government supporting degree and dairy industry cluster.

The development of the new economy geography school injected new energy to industry cluster theory. Baldwin (1999) believes that the government powerful protecting measures will help to increase local industry profit, attract capital, accelerate the capital accumulation, and promote the industry cluster. The study by Lanaspa et al., found the district where the government with higher efficiency is more attractive for the industry cluster by research. Dairy industry cluster can't get rid of the government's role. The local economic development protecting measures taken by government will be more helpful for the dairy industry cluster.

Hypothesis 6: There are positive correlation between economic basis and dairy industry cluster.

Dairy industry is part of the national economy. The district economic development level will affect this district industry development. In better economic based, and more developed districts, there will be better infrastructures, better investment development

environment, perfect rules guarantee, higher consuming level, and easier for the dairy industry clustering. In the poor economy based, and undeveloped districts, the infrastructure, investment environment and rules guarantee situations are worse, and it is not helpful for the dairy industry clustering. Therefore, the paper put forward the hypothesis that there are positive correlation between economic basis and dairy industry cluster.

Hypothesis 7: The regional industry profitability has positive influence on dairy industry cluster.

Represented by August Losch, the location theory of economic school believes that the internal competition decreases the cost and increases the demand. The concentrated demands will promote the manufacturing concentration. Industry clustered enterprises efficiency will be improved forced by the fierce competition, and the cost will be decreased. In this way it will meet the various marketing demands, and gain higher profit. Therefore, industry cluster region's profitability is stronger than others. Regional industry profitability will be very attractive for the industrial enterprises. The differences of the regional industry profitability will influence the dairy industry cluster.

Moreover, industry cluster theory considers that public transport infrastructure, and degrees of opening up to the world are the important factors of influencing industry cluster. But after experiencing a fast period of China transport infrastructure construction, it has basically formed the grades roadway, railway, high-speed rail, airlines and others multi-dimensional transportation system. Dairy industry, which is different from capital-intensive large scaled manufacturers, has weaker demand for transportation, so the transportation infrastructure situation will not be considered as the factor of influencing dairy industry cluster. China dairy industry enterprises are mainly domestic enterprises and the foreign countries enterprises of joint-stock. It is seldom for the foreign investors to build a factory to produce. And it has little influence on dairy industry cluster. Therefore, we don't take foreign investors into consideration in the influence factors of industry cluster.

Variables Selection

- Industry cluster degree: it uses the location quotient LQ to measure China dairy industry cluster degree. LQ, which was put forward by Haggett, is used to measure a certain region factors' space distribution situation. By calculating China different provincial dairy industry LQ, we could conclude that the dairy industry is relatively concentrated in which provinces. The accumulating formula is:
- $$LQ_{ij} = \frac{m_{ij} / gdp_{ij}}{M_j / GDP_j} \quad (1)$$
- m_{ij} refers to the jth year's dairy industry products sales turnover in the i province. The gdp_{ij} refers to the jth year's dairy industry products gross products in the i province. M_j is the jth year's dairy industry products sales turnover all over China. The larger is the LQ, the higher degree of dairy industry cluster will be. If LQ is more than one, we could believe that the dairy industry is highly clustered in this province.
- Resources endowment namely the availability of raw materials: the milk yield of the province could be the symbol to measure the dairy industry raw material availability. It will reflect the dairy industry resources endowment situation of every province. Represented in MY (milk yield).
- Labor force cost: the dairy industry employees' average wages is the best symbol of dairy industry labor force cost. But there are no unified standards of the employees' average wages in the dairy statistics annual, so the average wages of urban workers in every province institution will reflect dairy industry labor force cost. Represented in AWUW(average wages of urban workers)
- The current scale level: generally speaking, the larger is the industry assets scale, the

more output it will have. As it is difficult to get the existing assets statistics of every province dairy industry, and the relationship between assets scales and output ability, the paper selects the dairy products value to reflect the existing scale level of every province. Represented in DPV (dairy products value).

- The government support: the government revenue rate of GDP could reflect the government supporting degree in local marketing protection and support. The higher is the rate, more motivated government will be to support related industry development in this province. Represented by GRR (government revenue rate).
- The economic basis: there are many indexes of the region's economic basis. We usually use Per Capita GDP, gross GDP, government revenue etc. As the higher is the Per Capita GDP, the stronger the consuming capacity will be, the Per Capita GDP reflects the region's existing economic developed level. Represented in PCGDP (per capita GDP).
- Regional Industry profitability: the profit of dairy industry of every province could reflect the dairy industry profitability of every province. The higher is the profit rate, the stronger the profitability will be. We use the dairy industry profit rate divided by products sales revenue approximately reflects profitability. Represented by PM (profit margin).

Table 1. The influence factors of dairy industry cluster

Influence factors	Measuring index	Symbols abbreviation	Forecast Positive or Negative
Resource endowment	Milk Yield	MY	Positive
Labor force cost	Average wages of urban workers	AWUW	Negative
External scaled economy	Dairy products sales value	DPV	Positive
Government supporting degree	Government revenue rate	GRR	Positive
Economic basis	Per Capita GDP	PCGDP	Positive
Regional industry profitability	Profit Margin	PM	Positive

Basic Models Setting

- Based on the mentioned above hypothesis and related measuring index selecting, we set the liner model as below:

$$LQ_{ij} = \beta_0 + \beta_1 MY_{ij} + \beta_2 AWUM_{ij} + \beta_3 DPV_{ij} + \beta_4 GRR_{ij} + \beta_5 PCGDP_{ij} + \beta_6 PM_{ij} + \varepsilon_m \quad (2)$$

β_0 is constant; β_k is regression coefficient, k =1,2…6; i=1,2,3…29 represents the 29 provinces of China; j=1,2,3,4,5 represents the five years panel data from 2009 to 2103, ε is stochastic error.

- When selecting the samples, the study selects total 29 provinces without Hong Kong, Macau, Hainan, and Tibetan, as there is no neighbor of Hainan province, lack of some statistics of Tibetan. For the sake of the statistics availability and integrity, the statistics are mainly from 2010-2014 *Statistical Yearbook of China*, *Milk Yearbook of Chin,* and *China National Bureau* of *Statistics* official website. The data analysis is mainly done by advances spatial econometric scheme.

3. Results and discussions

Spatial Econometric Model Setting

In the traditional statistics theory, suppose the observation values are independent from each other, reviewing the spatial statistics, spatial econometric theory thinks that there are a few independent observation values. There is spatial interaction among observation values, namely there is spatial dependency and spatial autocorrelation among the regional economic geography statistics. Based on spatial econometric model, dairy industry cluster influence factors spatial econometric analysis should test if there is spatial autocorrelation in dependent variable. If there is, we should construct spatial autoregressive model and error models to spatial econometric estimation test of dairy industry influence factors.

(1) Dairy industry cluster spatial autocorrelation inspection

Global Spatial Autocorrelation will generally depict the provincial dairy industry distribution from provincial space. It is the important way to inspect if the industry cluster degree is high, and if the neighbor space point cluster is relevant. This paper inspects there is spatial relation among dairy industry clusters by using spatial autocorrelation index Moran's I. The formula is:

$$Moran'sI = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}(LQ_i - \overline{LQ})(LQ_j - \overline{LQ})}{S^2 \sum_{i=1}^{n}\sum_{j=1}^{n} W_{ij}} \quad (3)$$

And

$S^2 = \frac{1}{n}\sum_{i=1}^{n}(LQ_i - \overline{LQ})$; $\overline{LQ} = \frac{1}{n}\sum_{i=1}^{n} LQ_i$; LQ_i refers to the dairy industry cluster LQ coefficient of prince i, i=1,2,3…29; W_{ij} is nearby spatial weight matrix in binary, and it could recognize the neighborhood relation among different spaces. This paper uses weight matrix set in the neighborhood distance, namely if two provinces are neighbors, the W_{ij} is 1, if not W_{ij} is 0. The formula is as below:

$$W_{ij} = \begin{cases} 0 & \text{0 means province i and province j are not neighbors} \\ 1 & \text{1 means province i and province j are neighbors} \\ 0 & \text{0 means i=j} \end{cases}$$

And i=1,2,3…29,j=1,2,3…29, when i=j, Wij is 0, namely weight matrix diagonal line element is 0, $Moran'sI$ index is between -1 and 1. If $Moran'sI$ index is more than 0, there is positive spatial autocorrelation in dairy industry clusters. If $Moran'sI$ index is less than 0, there is negative spatial autocorrelation in dairy industry clusters. By drawing spatial correlation coefficient $Moran'sI$ scatter plot, China 29 provinces dairy industry cluster distribution is divided into four quadrants spatial dependency. The first quadrant HH (high-high), high dairy industry cluster degree province is surrounded by high cluster degree provinces; the second quadrant LH (low-high), the low dairy industry cluster degree province is surrounded by high cluster degree provinces; the third quadrant LL (low-low) the low dairy industry cluster province is surrounded by low cluster degree provinces; the fourth quadrant HL (high-low) the high dairy industry cluster province is surrounded by low cluster degree provinces.

According to $Moran'sI$, we can use the Normal Distribution Assumption to inspect if there is spatial autocorrelation among the 29 provinces. That is to calculate the $Moran'sI$ standard Z value under the condition of normal distribution assumption. The formula is as below:

$$Z(\mathrm{d}) = \frac{Moran'sI - E(Moran'sI)}{\sqrt{\mathrm{var}(Moran'sI)}} \quad (4)$$

The expected value

$$E(Moran'sI) = -\frac{1}{n-1}, \quad (5)$$

Variance

$$\mathrm{var}(Moran'sI) = \frac{n^2 W_1 + nW_2 + 3W_0^2}{W_0^2(n^2-1)} - E(Moran'sI) \quad (6)$$

If Z values are all more than the boundary 1.96 or 1.65 with the confidence level 0.01 or 0.05, there will be positive correlation in China dairy industry cluster, and there is significant spatial dependency.

(2). Spatial econometric model

There are two reasons for causing spatial autocorrelation. One is there is objective relation in neighborhood regions. The other is there is space error in selecting sample statistics, reflected in spatial autoregressive model error and dependent variables lags. Therefore, there are two spatial

econometric models: Spatial Lag Model, SLM, and Spatial Error Model, SEM.

- Spatial Lag Model, SLM

SLM is mainly used the province dairy industry cluster influence factors' influencing situation on dairy industry of neighborhood provinces, namely studying if there are spillovers of this variable in provinces. The expression is:

$$Y = \rho WY + X\beta + \varepsilon \quad (7)$$

Y is dependent variable, X is independent variable matrix, W is spatial weight matrix, WY is spatial autoregressive dependent variable, ρ is spatial auto-regression coefficient, ε is constant error. According to the above formula, dairy industry cluster influence factors spatial regression model is set as below:

$$LQ = \rho(E_T \otimes W)LQ + \beta_1 MY + \beta_2 AWUM + \beta_3 DPV + \beta_4 GRR + \beta_5 PCGDP + \beta_6 PM + \varepsilon \quad (8)$$

$(E_T \otimes W)$ is the kronecker products of matrix, E_T is T-order Matrices , W is spatial weight matrix, the parameter $\beta_i, i = 1,2...6$ reflects the dairy industry cluster influence factors' influence on cluster degree. Spatial regression variable $(E_T \otimes W)LQ_{ij}$ reflects the spatial distance influence on regional dairy industry cluster, and it is an endogenous variable.

- Spatial Error Model, SEM

There is the difference among different regions, we should use spatial error model to estimate. The formula of spatial error model is:

$$Y = X\beta + \varepsilon \quad (9)$$

$$\varepsilon = \lambda W\varepsilon + \mu = (E_n - \lambda W)^{-1}\mu \quad (10)$$

λ is spatial coefficient to measure spatial dependency, ε is random error, μ represents the random error vector matched the normal distributions, E_n is n-order matrix. According to the expression above, dairy industry cluster influence factor spatial error model set as below:

$$LQ = \beta_1 MY + \beta_2 AWUM + \beta_3 DPV + \beta_4 GRR + \beta_5 PCGDP + \beta_6 PM + \varepsilon + \mu \quad (11)$$

Parameter $\beta_i, i = 1,2...6$ reflects all the dairy industry cluster influence factors' influence on cluster degree, μ represents the random error vector matched the normal distributions, $\varepsilon = (E_n - \lambda W)^{-1}\mu$ 。

In the spatial econometric model, the estimate of traditional least square estimate variable coefficient is biased or even not effective. Spatial econometric model estimation usually uses the general least square estimate or maximum likelihood estimation. As for spatial lag and spatial error model, we adopt the method of maximum likelihood estimation.

Spatial Lag Model or Spatial Error Model, when we judge which model is more suitable to choose; generally we choose it according to the goodness-of-fit R2. The higher is R2, the higher goodness-of-fit will be, the better the fitting effect will be. What's more, we can choose according to Log likelihood (LogL), Likelihood Ratio (LR), Akaike information criterion (AIC), Schwartz criterion (SC).

The larger is LogL, the smaller AIC, SC are, the higher is the goodness-of-fit; the smaller LogL is, the larger AIC, SC are, the lower is the goodness-of-fit. We could taking R2, LogL, AIC, SC vales into consideration when we choose spatial lag model and spatial error model.

Discussions of empirical results analysis

(1)Dairy industry cluster spatial autocorrelation inspection.

According to LQ expression (1), China provinces dairy industry LQ coefficients shows in Table 2. Reviewing from regions, China dairy industry heavy cluster is mainly located in north China, northeast China, and northwest China. While south China, southwest China, and east China is with light cluster. Reviewing from time levels, from 2009 to 2013,dairy industry cluster degree generally increased fast in northwest China, especially in Ningxia and Shaanxi Province; but the traditional dairy industry provinces Inner Mongolia and Heilongjiang industry cluster degree decreased significantly from 2009 to 2013. In 2009, the top six provinces with heavy dairy industry cluster degrees are: Inner Mongolia, Heilongjiang, Ningxia, Shaanxi, Shanghai, and Hebei province. In 2013, the top six provinces with heavy dairy industry cluster degrees are: Heilongjiang, Ningxia, Inner Mongolia, Shaanxi, Hebei, and Shanghai. There is no change of the top six provinces, but there is a sequence change. Inner Mongolia dairy industry cluster degree decreased to the third rank from the first place; Heilongjiang increased to the first rank from the second; Ningxia increased from the third to the second; Shanghai decreased to the six; Hebei increased to the fifth from the sixth. However, reviewing from the cluster degree values, it is decreasing in Heilongjiang, Inner Mongolia and Shanghai; it is increasing in Ningxia, Shaanxi and Hebei province. Generally reviewing from 2009 to 2013, China most provinces dairy industry cluster degree were increasing, a few provinces cluster degree decreased differently. The traditional dairy industry clustered regions continued dairy industry, and there is no obvious new industry cluster. But the cluster degree absolute value is increasing and decreasing. The quick increasing of Ningxia, Shaanxi and Hebei dairy industry cluster degree injected new energy to China dairy industry development.

Table 2. Every province dairy industry LQ coefficient

	Province	2009	2010	2011	2012	2013
North	Beijing	1.078415	1.034069	1.069968	1.077882	1.102943
	Tianjin	0.502999	0.618738	0.865858	0.713553	1.365149
	Hebei	1.503711	1.567345	1.656046	1.74751	1.785628
	Shanxi	0.750314	0.937223	0.783025	0.784102	0.758432
	Inner Mongolia	7.202618	6.522545	5.909151	4.824749	4.710306
Northeast	Liaoning	1.243544	1.19077	1.194555	1.073878	1.059847
	Jilin	0.246448	0.340034	0.479164	0.484211	0.4691
	Heilongjiang	6.83371	6.664661	6.219887	5.221178	5.184054
East	Shanghai	1.510664	1.565186	1.503468	1.720428	1.478702
	Jiangsu	0.218232	0.217845	0.260504	0.352624	0.369237
	Zhejiang	0.22644	0.188895	0.230163	0.294521	0.284968
	Anhui	0.751748	0.81895	0.733367	0.839488	0.803823
	Fujian	0.234334	0.207023	0.109558	0.130401	0.134807
	Jiangxi	0.515701	0.474673	0.410789	0.47635	0.493416
	Shandong	1.037763	1.075496	1.072848	1.296846	1.189423
South	Henan	0.497506	0.546818	0.711102	0.790279	0.836593
	Hubei	0.463306	0.469348	0.470466	0.466516	0.500674
	Hunan	0.743311	0.524612	0.520196	0.365889	0.402325
	Guangdong	0.50279	0.544303	0.565706	0.597154	0.62086
	Guangxi	0.199281	0.304914	0.343655	0.370322	0.433459
Southwest	Chongqing	0.309863	0.475355	0.488744	0.398242	0.545348
	Sichuan	0.384578	0.502431	0.538048	0.512992	0.540088
	Guizhou	0.367003	0.448483	0.510066	0.187972	0.172577
	Yunnan	0.429736	0.445402	0.465339	0.634735	0.621552
Northwest	Shaanxi	1.744979	1.736334	1.886317	2.030003	2.144147

	Gansu	0.42323	0.529859	0.476545	0.805404	0.839343
	Qinghai	0.820136	1.376558	0.604174	0.906551	0.822348
	Ningxia	2.147102	1.772315	2.263251	4.400071	4.8391
	Xinjiang	1.018264	1.113483	0.975052	1.041966	0.899895
	Province	2009	2010	2011	2012	2013
North	Beijing	1.078415	1.034069	1.069968	1.077882	1.102943
	Tianjin	0.502999	0.618738	0.865858	0.713553	1.365149
	Hebei	1.503711	1.567345	1.656046	1.74751	1.785628
	Shanxi	0.750314	0.937223	0.783025	0.784102	0.758432
	Inner Mongolia	7.202618	6.522545	5.909151	4.824749	4.710306
Northeast	Liaoning	1.243544	1.19077	1.194555	1.073878	1.059847
	Jilin	0.246448	0.340034	0.479164	0.484211	0.4691
	Heilongjiang	6.83371	6.664661	6.219887	5.221178	5.184054
East	Shanghai	1.510664	1.565186	1.503468	1.720428	1.478702
	Jiangsu	0.218232	0.217845	0.260504	0.352624	0.369237
	Zhejiang	0.22644	0.188895	0.230163	0.294521	0.284968
	Anhui	0.751748	0.81895	0.733367	0.839488	0.803823
	Fujian	0.234334	0.207023	0.109558	0.130401	0.134807
	Jiangxi	0.515701	0.474673	0.410789	0.47635	0.493416
	Shandong	1.037763	1.075496	1.072848	1.296846	1.189423
South	Henan	0.497506	0.546818	0.711102	0.790279	0.836593
	Hubei	0.463306	0.469348	0.470466	0.466516	0.500674
	Hunan	0.743311	0.524612	0.520196	0.365889	0.402325
	Guangdong	0.50279	0.544303	0.565706	0.597154	0.62086
	Guangxi	0.199281	0.304914	0.343655	0.370322	0.433459
Southwest	Chongqing	0.309863	0.475355	0.488744	0.398242	0.545348
	Sichuan	0.384578	0.502431	0.538048	0.512992	0.540088
	Guizhou	0.367003	0.448483	0.510066	0.187972	0.172577
	Yunnan	0.429736	0.445402	0.465339	0.634735	0.621552
Northwest	Shaanxi	1.744979	1.736334	1.886317	2.030003	2.144147
	Gansu	0.42323	0.529859	0.476545	0.805404	0.839343
	Qinghai	0.820136	1.376558	0.604174	0.906551	0.822348
	Ningxia	2.147102	1.772315	2.263251	4.400071	4.8391
	Xinjiang	1.018264	1.113483	0.975052	1.041966	0.899895

We can calculate China 29 provinces general $Moran'sI$, E(I), Z, sd(I), and P value from 2009 to 2013 by using formulas(3) (4) (5) (6), as shows in Table 3. From 2009 to 2013, the $Moran'sI$ P values are all less than 0.05, Z value is all more than the boundary 1.96 with the confidence level 0.05. $Moran'sI$ value increased to 0.336 in 2013 from 0.273 in 2009. It shows that there are obviously positive spatial spillovers in China 29 provinces dairy industry cluster, and there is spatial dependency among the 29 provinces. The dairy industry spatial distribution is not independent; instead, the provinces with higher cluster degrees are neighbors, the provinces with lower cluster degrees are neighbors. The spatial cluster is very obvious. To further analyze dairy industry spatial cluster characteristics, the paper made the $Moran'sI$ scatterplot, shows as Graph 1. This $Moran'sI$ and Table 4 shows most of our provinces are distributed in the first quadrant and the third quadrant. The first quadrant represents the higher cluster degree province are surrounded by other higher cluster degree provinces, including Ningxia, Shaanxi, Inner Mongolia, Heilongjiang, Hebei, Tianjin total six provinces, which accounts for 20.6% of the whole provinces, mainly locates in northwest and north China.

Table 3. China 29 provinces dairy industry cluster $Moran'sI$ index value

Year	Morna's I	E(I)	sd(I)	z	P-value
2009	0.273	-0.036	0.101	3.067	0.002
2010	0.270	-0.036	0.101	3.028	0.002
2011	0.308	-0.036	0.103	3.334	0.001
2012	0.331	-0.036	0.113	3.254	0.001
2013	0.336	-0.036	0.113	3.297	0.001

Note: Hainan, who doesn't have neighbors, was not studied; Hong Kong, Macau, Taiwan, and Tibet

Autonomous Region, whose data is not available, were not studied.

The second quadrant refers to the lower cluster degrees provinces are surrounded by higher cluster degree provinces, including Beijing, Shaanxi, Liaoning, Jilin, Gansu total five provinces, mainly locates in north and northeast China; The third quadrant refers to the lower cluster degrees provinces are surrounded by lower cluster degree provinces, including Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Qinghai, and Xinjiang total 17 provinces, accounts for 58.62% of the whole provinces, mainly locates in east and south and southwest China; The fourth quadrant refers to the higher cluster degrees provinces are surrounded by lower cluster degree provinces, only with Shanghai in the fourth quadrant, accounts for 3.45% of the whole provinces. The above analysis shows: one is that there is spatial cluster phenomenon in China dairy industry cluster; the other is that there is spatial dependency and heterogeneity in industry cluster, it proves the Hypothesis 1. Therefore, we should analyze dairy industry influence factors from spatial econometric perspective.

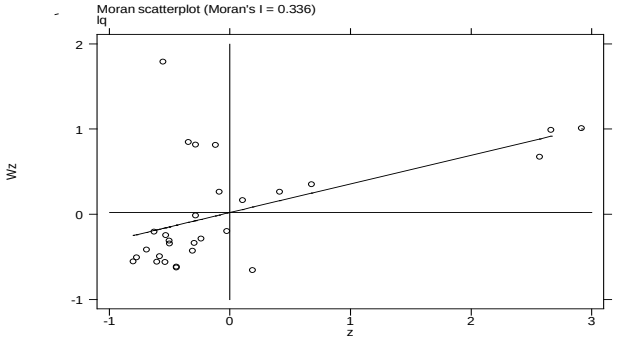

Figure 1. China dairy industry cluster $Moran'sI$ scatter plot

Dairy industry spatial econometric model estimation

The spatial correlation inspection result has showed there is obvious spatial autocorrelation in China 29 provinces dairy industry cluster. Traditional least square regression estimation result is biased. Therefore, the paper adopts the SLM and SEM under the fixed effects and random effects, and estimates and inspects the dairy industry cluster influence factors with China 29 provinces spatial panel data basis. It will choose the best model and explain according the inspection results. To compare easily, first it makes the ordinary least square regression. The results show in Table 5. The least square estimation general goodness-of-fit R^2 is 0.7143. Resources endowment, external economy, government policies environment regression coefficient are positive, and the distribution passed the 5% and 1% confidence level inspection. It shows that there is positive correlation between these three variables and dairy industry cluster. While the economy basis coefficient is negative and passed the 5% confidence level inspection, this is possibly because of the neglecting of spatial autocorrelation. Therefore, it needs spatial econometric model to estimate and inspect dairy industry cluster.

Table 4. Different quadrants provinces distribution

Quadrant No.	Spatial Relevant ways	Province
Quadrant I	HH(high-high)	Ningxia, Shaanxi, Inner Mongolia, Heilongjiang, Hebei, Tianjin
Quadrant II	LH(low-high)	Beijing, Shanxi, Liaoning, Jilin, Gansu
Quadrant III	LL(low-low)	Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangdong, Guangxi, Chongqing, Sichuan, Guizhou, Yunnan, Qinghai, and Xinjiang
Quadrant Ⅳ	HL(high-low)	Shanghai

The parameter estimation of fixed effects and random effects spatial lag model and spatial error model used maximum likelihood

estimation method. The result shows as in Table 6. In the spatial error model result shows, the P value of CHI^2 in the Hausman testing is 0.9613. It rejects the original hypothesis and adopts the fixed effects. Meanwhile, the AIC and SC value of fixed effects are much more less than random effects.

Table 5. Least Square Estimation OLS

Variable	Coefficient ρ	Standard Deviation σ	t statistic value	P value
MY	3.39e-07**	6.78e-08	5.00	0.000
AWUW	6.73e-06	0.0000144	0.47	0.642
DPV	0.0073747***	0.0017541	4.20	0.000
GRR	10.57208***	4.029894	2.62	0.010
PCGDP	-0.0000173**	7.10e-06	-2.44	0.016
PM	-1.211589	2.304065	-0.53	0.600
cons	-0.4989081	0.3605385	-1.38	0.169
R2	0.7143			
AIC	352.927			
SC	373.7641			

Note: ***, **, * represents respectively 1%, 5% and 10% confidence level significant test 。

Therefore, fixed effects spatial error model is better than the random effects spatial error model. But the fixed and random effect of spatial lag model parameter λ doesn't pass the 10% confidence level significance testing. Therefore, the spatial error model is not goodness-of-fit for the whole model. The spatial lag model results shows the P value of CHI^2 in the Hausman testing is 0.973. It rejects the original hypothesis and adopts the fixed effects. Meanwhile, the AIC and SC value of fixed effects are much more less than random effects, and the R^2 of the fixed effects is 0.8062, and it is more than the R^2 of random effects. That is to say, the fixed effects spatial lag model is with the best goodness-of-fit., and the fixed effects spatial lag model parameter ρ passed the 10% confidence level significance testing. Compared with the above mentioned five spatial econometric models, the paper chooses the best goodness-of-fit fixed effects spatial lag model, and defines the parameter estimated economy.

Table 6. Dairy industry cluster spatial lag model and spatial error model estimation results

Variable	Spatial Error Model, SEM				Spatial Lag Model, SLM			
	Fixed effects	P value	Random effects	P value	Fixed effects	P value	Random effects	P value
MY	3.43e-07***	0.000	3.50e-07***	0.000	3.40e-07***	0.000	3.28e-07***	0.000
AWUW	8.76e-06	0.529	-8.86e-07	0.942	0.0000112	0.390	-5.90e-07	0.960
DPV	0.0075375***	0.000	0.0070135***	0.000	0.0075642***	0.000	0.0069709***	0.000
GRR	16.70575***	0.001	15.31339***	0.001	15.89509***	0.001	15.03425***	0.001
PCGDP	-0.0000356***	0.005	-0.0000219**	0.023	-0.0000373***	0.002	-0.000022**	0.017
PM	2.705108**	0.033	2.157213	0.119	2.629744**	0.036	2.232159*	0.102
cons			-0.677795*	0.086			-0.7964253**	0.041
σ^2	0.0908785***	0.000	0.1146391***	0.000	0.0909542***	0.000		
σ							0.1159495***	0.000
ρ					0.1147027*	0.101	0.1354406	0.155
λ	0.0496378	0.722	0.0914267	0.513				
R2	0.6979		0.7294		0.8062		0.7464	
LogL	-31.9151		-94.6815		-31.9479		-93.9300	
AIC	77.83026		207.363		76.89588		205.86	
SC	98.6674		234.1536		97.73302		232.6506	
Hausman	Prob>chi2 = 0.9613				Prob>chi2 = 0.9735			

Note: Same as Table 5.

Results Analysis

Hypothesis 1 is supported. The Table 6 results that fixed effects spatial lag model ρ passed the 10% confidence level significance testing. It means China dairy industry cluster is becoming gradually obvious, and there is spatial autocorrelation and spatial spillovers. The paper will analyze dairy industry cluster influence factors one by one.

Hypothesis 2 is supported. Fixed effects spatial lag model MY coefficient is 3.43e-7, and it passed the 1% confidence level significant testing. It has positive correlation with dairy industry cluster. It is one of the factors influencing dairy industry cluster. But due to the improvement of transportation infrastructure and quality keeping and fresh-keeping technology, raw milk transporting and cool-keeping are no the key problems that dairy industry enterprises are concerning.

Hypothesis 2 is rejected. Fixed effects spatial lag model labor cost AWUW coefficient estimation doesn't pass the 10% confidence level significance testing. That is to say the labor cost will not significantly influence China dairy industry cluster. It is incompatible with the materials endowment theory of Hypothesis 1. The reasons are: one is dairy industry clustered provinces Heilongjiang, Inner Mongolia, Shaanxi, Ningxia natural conditions, location, infrastructure aspects are behind the east coastal regions. Thus it will somehow offset the labor cost advantage. The other one is the existing of snowballing effects. Dairy industry once clustered in the provinces, it will develop very well and self-improved well. And it will further improve cluster, and finally it forms the snowballing effects, and weakens the cheap labor cost's influence on dairy industry cluster.

Hypothesis 4 is supported. Fixed effects spatial lag model existing scales level DPV coefficient is 7.5375e-3, it passed 1% confidence level significance testing, and it has positive correlation with dairy industry cluster. This means the dairy industry external economic effect truly exists in Hypothesis 4. When the dairy industry is cluster in a province, it will promote the dairy cluster with the whole scale increasing and general cost decreasing. Namely when the existing scale level increase one by unit, dairy industry cluster degree will increase 7.537e-3 by unit.

Hypothesis 5 is supported. Fixed effects spatial lag model government supporting degree GRR coefficient is 15.89509, and it passed 1% confidence level significance testing, and it has obvious positive correlation with dairy industry cluster. It proves that the government supporting degree provides good developing environment for the whole industry, and it will further exist industry cluster in the region hypothesis. Meanwhile it means government is playing important role in industry developing and industry clustering. When the coefficient is 15.89509, it is namely when government revenue percentage for GDP increase by 1%, cluster degree coefficient will increase by 0.1589509.

Hypothesis 6 is incompatible. Fixed effects spatial lag model economic basis PCGDP coefficient is -3.73e-5, and it passed 1% confidence level significance testing. It has negative correlation with dairy industry cluster. So it is incompatible with hypothesis 6. That is to say provincial economic development will not positively influence dairy industry cluster. On the contrary, it is negative correlated. When Per Capita GDP increases 1 by unit, cluster degree will decrease 3.73e-5 by unit. The reason is that the Per Capita GDP of east China economy developed provinces is much higher than the middle and west China. However, factory construction, production managing, enterprise managing, employees wages costs are also much higher than middle and west China. Meanwhile, we have to think about the raw milk availability, that most of the dairy industry enterprises are located mainly in the undeveloped provinces like Inner Mongolia, Heilongjiang, Ningxia, Shaanxi province. Therefore, there comes the result of negative correlation between economic development and diary industry cluster.

Hypothesis 7 is supported. Fixed effects spatial lag model regional profitability PM coefficient is 2.629744, and it passed the 5% confidence level significance testing. It has positive correlation with dairy industry cluster. It proves Hypothesis 7 that the higher is the dairy industry profit margin rate, the more attractive for the enterprise it will be.

4. Conclusions

The paper finally choose fixed effects spatial lag model to explain China diary industry cluster influence factors by comparing ordinary least square estimation (OLS), and constructing fixed effects and random effects spatial lag model(SLM) and spatial error model (SEM), and come to the conclusion below:

There is positive correlation among China dairy industry cluster. There is strong spatial dependency and positive spatial spillovers in different provinces, namely, the neighborhood provinces' dairy industry development will promote this province dairy industry development and cluster. Therefore, it will enhance China dairy industry develop in a healthy and fast way by communicating and cooperating more with neighborhood provinces, and promoting regions rational flow of talents and capitals

The results of the most of industry cluster studies with time dimensions are similar. There is significant correlation among the resources endowment, external scope economy, industry profitability and China dairy industry cluster. According to the self-raw milk availability characteristics, and the existing scale level, different provinces takes different diary industry developing strategies. It is playing significant practical role in promoting diary industry cluster development and improving industry branded enterprise' competition. There is significant positive correlation between government supporting degree and dairy industry cluster. Government maintaining good economic running environment is the catalyst and the guarantee of the dairy industry cluster. Based on every province's different dairy industry developing levels, the province will push dairy industry further development with government and marketing joint efforts. Different from the hypothesis, labor force is no the significant factor of influencing dairy industry cluster. The dairy industry clustered provinces are not getting benefit from the labor force cost. Incompatible with the hypothesis, there is negative correlation between the economic basis and dairy industry cluster. Compared with developed provinces, dairy industry tends to cluster in the undeveloped provinces. Middle and west China should seize this opportunity to promote it's dairy industry cluster, and activate the economic development.

While, there is broad space to improve the paper. On the one hand, the diary industry cluster influence factors spatial econometric model constructed in the paper could be expanded, for instance, technology, chance factors could be introduced into the model to be further observed and studied. On the other hand, the penal data yeas span is limited in this paper, without enough consideration of time dimension characteristics, and the class weights setting without consideration of economic distance. If the paper uses long time span dynamic spatial panel econometric model, the conclusion will be more purposeful and persuasive.

5.References

Baldwin, R.E. (1999). Agglomeration and endogenous capital, *European economic review*, 43, 253-280.

Cheng, Xiaoping. (2012). China dairy industry spatial cluster and the influence on industry growing, *Journal of Wuhan University of Technology (Social Scientific Edition)*, 05, 688-693.

Easterly, W, Levine, R. (1998). Troubles with the neighbors: Africa's problem, Africa's opportunity, *Journal of African economics*, 7(1),120-142.

Fujita, M.A. (1988). Monopolistic competition model of spatial agglomeration: differentiated product approach, *Regional science and urban economics*, 18(1), 87-124.

Henderson, J.V. (1974). The sizes and types of cities, *American economic review*, 64(4), 640-656.

Hua, Junguo, Zhu, Xiangrong, Yin, Chenwen. (2007). The analysis of China dairy cluster situation and spatial distribution, *China rural economy*, 02, 49-54+80.

Lanaspa, L.F, Pueyo, F. & Sanz F. (2001). The public sector and core-peripthery models, Urban studies, 38,1639-1649.

Lesage, J P.A (1999). Spatial econometric examination of China's economic growth, *Geographic information sciences*, 5(2),143-153.

Li, Li. (2008). Industry investment location selecting and industry cluster—take Inner Mongolia Animal Products Industry as an example, *China rural economy*, 01,12-22.

Li, Li, An, Yufa. (2008). Regional production specialization and industry cluster — analysis of Inner Mongolia animal products industry cluster trends. *Agricultural economic issues* ,05,44-49.

Liu, Hedong. (2013). Empirical study of regional creation inside overflow, outside overflow, and spatial spillover effect, *Science research management*, 01, 28-36.

Niu, Honglei, Jiang, Keshen. (2011). The panel data analysis of China textile Industry cluster Distribution layout and influence factors, *Application of statistics and management*, 04, 571-584.

Ohlin, B. (1968). Interregional and International Trade. Cambridge: Harvard University Press, (89-92).

Ren, Yinghua, Xu, Ling, You, Wanhai. (2010) Financial cluster influence factors and application, *The journal of quantitative technical economics*, 05,104-115.

Sun, Qinggang, Guo, Ju'e, Shi, Bo. (2013). China Provincial Energy Intensity Spatial Spillover Analysis, *China's population, resources and environment*, 11, 137-143.

Tschoegl, Adrian E. (2000). International banking Centers, geography, and foreign banks, financial markets, institutions & instruments, 9 (1), 1-321.

Wang, Yongpei, Yuan, Pinghong. (2010). Wages difference, labor force flowing, and industry clustering—based on new economic geography explanation and empirical testing, *Finance and economics*,03,59-66.

Zhang, Xueliang. (2012). Does China communication infrastructure promote regional economy growth—the communication infrastructure spatial spillovers effect, *Social science of China*, 03,60-77; 206.

Zhao, Liangshi, Sun Caizhi, Zheng Defeng. (2014). China Inter-provincial water resource utilization efficiency and spatial spillover effect measure, *Geography journal*, 01,121-133.