Spatio-temporal Agglomeration and Coupling Effects of Inputs and Outputs in China’s Biomedical Industry

Yanchao Gao1*, Sunyun Qi1, Siyuan Chen2, Hua Gu3, Qifeng Zhang1, Meiying Gao1

1 Center for Medical Science Technology and Education Development, Hangzhou, Zhejiang 310002, China
2 Leuven Statistics Research Centre (LStat), KU Leuven, Leuven 3000, Belgium
3 Institute of Basic Medicine and Cancer, Chinese Academy of Sciences 310000, China

Abstract: Biomedical industry is one of the major development industries in the world and academics are interested in the inputs and outputs of biomedical industry. We used the inputs and outputs data of biomedical industry of 31 provinces and municipalities in China from 2009 to 2020 (except 2017). Using the Theil index and Moran’s I, the spatial distribution and agglomeration of the inputs and outputs of the biomedical industry among 31 provinces and municipalities were analyzed. To examine the coupling effects and possible relevant factors for the inputs and outputs of the biomedical industry, the coupling coordination degree (CCD) and spatial lag model (SLR) were used. The overall development of the biomedical industry in China's 31 provinces and municipalities was found to be relatively balanced, with a high concentration in Jiangsu, Shandong, Zhejiang, and other provinces. Moreover, there is a widening development disparity between provinces. Government investment and enterprise investment have a significant impact on the development of the biomedical industry, with enterprise R&D capital investment being the most influential factor. In the meantime, the development of the biomedical industry has a close relationship with the investment and development levels of neighboring provinces.

Keywords: Biomedical Industry, Inputs and Outputs, Spatio-temporal Agglomeration, Coupling Coordination Degree, Moran’s I, Spatial Lag Model

1. Introduction

Biomedical industry is regarded as a sunrise industry, and the average annual growth rate of the global biomedical industry, statistically speaking, will reach 7%, which is much larger than the average growth rate of global GDP [1]. Major countries such as the United States, the United Kingdom, Japan, Germany, etc., have prioritized the development of biomedical industry [2], with the biomedical industry in the United States accounting for approximately 17% of GDP. Nearly 27 provinces in China have released their 14th Five-Year Plans in 2021, and they all promote the biomedical industry as a key sector for national economic development [3-4]. Consequently, it is of vital importance to study the development of the biomedical industry.

Operating revenue and revenue from new products are among the most significant indicators for measuring the level of development in the biomedical industry [5-6], revealing the overall level of development in 31 provinces. Government investment and enterprise investment are closely related to operating revenue and revenue from new products because the biomedical industry has obvious policy-driven characteristics [7-8] and the enterprise R&D investment is the most influential factor driving the development of the biomedical industry [9-10].

2. Materials and methods

2.1 Data source

The data of operating revenue, government investment, enterprise investment, and revenue from new products of the biomedical industry in 31 provinces and municipalities are from the 2009-2021 China Statistical Yearbook on High Technology Industry, but the 2017 data are missing due to the fact that the 2018 China Statistical Yearbook on High Technology Industry has not been published. Hong Kong, Macao, and Taiwan are not included due to a lack of statistical information.

2.2 Methods

2.2.1 Theil index

The Theil index, proposed by econometrician Henri Theil in 1967, is a statistic used to measure economic inequality between regions. From 2009 to 2020, the Theil index was
used to measure the difference in operating revenue of the biomedical industry across 31 provinces and municipalities in China. The smaller the Theil index, the smaller the disparity between each province's biomedical industry development, and vice versa. The formula is:

\[ T = \frac{1}{N} \sum \left( \frac{X_i}{\mu X} \ln \left( \frac{X_i}{\mu X} \right) \right) \]

where \( T \) is the Theil index; \( x_i \) is the operating revenue of the \( i \)th province; and \( \mu x \) is the average of the operating revenue.

### 2.2.2 Coupling coordination degree

The degree of coupling \( C \) reflects the degree of interdependence of systems. The degree of coupling coordination \( D \) measures the association degree of coupling between systems and reflects the coordination quality. The coupling coordination degree model of two systems of biomedical industry development inputs \( U1 \) and outputs \( U2 \) is as follows:

\[ D = \sqrt{C \cdot T}, C = \frac{1}{2} \left( \frac{U1 + U2}{U1 + U2} \right), T = a U1 + b U2 \]

where \( U1 \) is the inputs index of biomedical industry development, calculated from the government investment and the enterprise investment; \( U2 \) is the outputs index of biomedical industry development, calculated from the operating revenue and the revenue from new product; \( T \) represents the comprehensive coordination index of two systems. \( a \) and \( b \) are undetermined coefficients, with the sum of 1. In this article, both \( a \) and \( b \) were assigned the value 0.5, indicating that inputs and outputs systems carry the same weight.

### 2.2.3 Moran's I and spatial lag model (SLR)

Moran’s I is a statistic to measure spatial autocorrelation, which can qualify how closely clustered different features are and test the spatial effect before SLR is carried out. Global Moran’s I takes values between -1 and 1, where -1 means the variable of interest is perfectly dispersed. 0 indicates the variable of interest is randomly dispersed, and 1 means the variable of interest is perfectly clustered together. The formula is as follows:

\[ I = \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} z_i z_j}{S_0 \sum_{i=1}^{n} z_i^2} ; S_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \]

where \( z_i \) represents the deviation of the variable of interest from its mean for the \( i \)th province; \( w_{ij} \) is the spatial weight between \( i \)th and \( j \)th province; \( n \) is the total number of provinces; \( S_0 \) is the sum of all \( w_{ij} \).

Spatial autoregressive model (SAR) is a spatial method to describe the relationship between independent variables \( X \) and dependent variable \( Y \) by involving the spatial effect. The general formula is as follows:

\[ Y = \rho WY + X\beta + \lambda Wu + \epsilon ; \epsilon \sim N(0, \delta^2 I_n) \]

where \( W \) is the spatial weighting matrix; the parameter \( \rho \) is the spatial lag coefficient; \( \beta \) is a vector of regression coefficient, the parameter \( \lambda \) is the spatial error coefficient; and \( \epsilon \) is a vector of residual. In this article, SLR model was conducted, which is one of the special cases of SAR without spatial error term \( \delta Wu \).

All analyses were performed using GeoDa software, version 14.0.

### 3. Results

#### 3.1 Theil index

Theil index of the biomedical industry's operating revenue in 31 provinces from 2009 to 2020 is between 0.4 and 0.5, close to 0, indicating a relatively balanced development, as shown in Table.1. and Figure 1. From 2009 to 2016, the Theil index shows an upward trend, which demonstrates an uneven development trend. From 2016 to 2020, there is an obvious decreasing trend, which means the imbalance development has been improved.

**Table.1.** 2009-2020 Theil index of biomedical industry's operating revenue

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Theil index</td>
<td>0.73</td>
<td>0.69</td>
<td>0.65</td>
<td>0.62</td>
<td>0.59</td>
<td>0.57</td>
<td>0.55</td>
<td>0.53</td>
<td>0.51</td>
<td>0.49</td>
<td>0.47</td>
<td>0.45</td>
</tr>
</tbody>
</table>

**Figure 1.** 2009-2020 Theil index of biomedical industry's operating revenue

#### 3.2 Moran’s I

The biomedical industry’s operating revenue, government investment, and enterprise investment of 31 provinces in 2009, 2012, 2016, and 2020 were chosen as variables of interest to calculate the Moran’s I. Figure 2 exhibits the LISA cluster maps of government investment (a column), enterprise investment (b column), and operating revenue (c column) in 2009, 2012, 2016, and 2020 for the 31 provinces’ biomedical industry, based on the values of local Moran’s I. The values of local Moran’s I based on the biomedical industry’s operating revenue in 2009, 2012, 2016, and 2020 are 0.233, 0.244, 0.227, and 0.278, an overall upward trend, which indicates the biomedical industry has a trend to agglomerate. According to the c
can be noticed that the FEMS 2023 SHS Web of Conferences 169, 01032 (2023) https://doi.org/10.1051/shsconf/202316901032 FEMS 2023

column maps in Figure 2. Shandong, Jiangsu, Shanghai, Anhui, and Zhejiang provinces form a highly clustered area, and the operating revenue of the biomedical industry in Sichuan province is significantly higher than that in the neighboring provinces. Shanghai and Anhui shift over time from low output value to high output value in biomedical industry. In terms of government investment, the local Moran’s I for 2009, 2012, 2016, and 2020 are 0.075, 0.089, 0.164, and 0.047, respectively, and the value is larger in 2016. The map of 2016 government investment in Figure 2 shows the formation of a high-level government investment area that includes Shandong, Jiangsu, and Anhui. The local Moran’s I values of enterprise investment are 0.377, 0.269, 0.278, and 0.289, indicating an obvious clustering effect. Shandong, Jiangsu, Anhui, Shanghai, and Zhejiang comprise a region for high-level enterprise investment, according to the b column maps in Figure 2. It can be noticed that the enterprise investment is aligned with the evolution trend of operating revenue.

3.3 Coupling coordination degree

The government investment and the enterprise investment were used to compute the index for the 31 provinces’ biomedical industry, while the operating revenue and the revenue from new products were used to calculate the output index. Table 2 displays the coupling degree C, the coupling coordination degree D, and the comprehensive coordination index T. According to the values, three categories can be classified:

< 0.2, “low inputs-low outputs” 17 provinces: Shanxi, Inner Mongolia, Liaoning, Heilongjiang, Fujian, Jiangxi, Guanxi, Hainan, Chongqing, Guizhou, Yunnan, Tibet, Shaanxi, Gansu, Qinghai, Ningxia, and Xinjiang.

0.21-0.49, “with inputs-with outputs” 10 provinces: Beijing, Tianjin, Hebei, Jilin, Shanghai, Anhui, Henan, Hubei, Hunan, Sichuan, in which Beijing, Anhui, and Sichuan inputs and outputs growth is obvious, coupled with a significant increase in coordination.

> 0.5, “high inputs-high outputs” 4 provinces: Jiangsu, Zhejiang, Shandong, and Guangdong.

<table>
<thead>
<tr>
<th>Year</th>
<th>C</th>
<th>T</th>
<th>D</th>
<th>C</th>
<th>T</th>
<th>D</th>
<th>C</th>
<th>T</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2012</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2016</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>2020</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 2. Indexes in the coupling coordination degree model for biomedical industry’s inputs and outputs in 31 provinces.

Figure 2. LISA cluster maps of government investment (a column), enterprise investment (b column), and operating revenue (c column) for the biomedical industry in the 31 provinces in 2009, 2012, 2016, and 2020 (from top to bottom).


3.4 Spatial lag model

Table 3 shows the fitting results of a simple linear regression model for operating revenue. According to the value of $R^2$, the independent variables government investment and enterprise investment explain approximately 94% of the change in the operating revenue and both have significantly positive effect on the operating revenue. The effect of enterprise investment on operating revenue is much larger than that of government investment.

Table 3. Linear regression model results for operating revenue

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std.Error</th>
<th>t-Stat</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONSTAN T</td>
<td>181.365</td>
<td>42.8071</td>
<td>4.2367</td>
<td>0.00019</td>
</tr>
<tr>
<td>Government investment</td>
<td>0.0018943</td>
<td>0.0001579</td>
<td>11.991</td>
<td>0.00000</td>
</tr>
<tr>
<td>Enterprise investment</td>
<td>0.0227176</td>
<td>0.00313</td>
<td>3.2043</td>
<td>0.00013</td>
</tr>
</tbody>
</table>

R2=0.944596, Adjusted R2=0.940121, F=264.262**, Log likelihood=-226.261, Akaike info criterion (AIC)=458.522, Schwarz criterion (SC)=463.101

Furthermore, the SLR model was also used to fit the operating revenue, and the results are shown in Table 4. Compared with the simple linear regression model, the SLR model fits better, with a larger log-likelihood value, smaller AIC, and smaller SC. The likelihood ratio test was performed to diagnose for spatial dependence and gave a test statistic of 5.80 with a P-value less than 0.05, which means there is a spatial feedback effect. Neighboring provinces can affect the operating revenue of the local biomedical industry. The government investment and the enterprise investment have a significantly positive effect on the operating revenue. The spatial feedback effect coefficient is 0.257, with a P-value less than 0.05, indicating that the operating revenue of the biomedical industry in a certain province has mutual influence with neighboring areas via spatial feedback effect and that there may be spatial spillover effects.
4. Conclusion

Theil index shows that the biomedical industry’s overall growth in 31 provinces is evenly distributed. However, there is a trend of increasing development gaps between the provinces. In 2016, policies such as the Guidance from the General Office of the State Council on Promoting the Healthy Development of the Pharmaceutical Industry increased the significance of biomedical industry development in Chinese provinces. It results in a narrowing of the gap between the provinces’ biomedical industry development from 2016 to 2020. According to the values of local Moran’s I, Shandong, Jiangsu, Shanghai, Anhui, and Zhejiang are high concentration areas for the biomedical industry. In terms of enterprise investment, Shandong, Jiangsu, Anhui, Shanghai, and Zhejiang are high-level input areas. Furthermore, the consistency of the evolution trend between the enterprise investment and the operating revenue indicates a certain coupling. Based on the coupling coordination degree model of the inputs and outputs of biomedical industry in 31 provinces, the overall input-output relationship can be divided into three levels. Only Jiangsu, Zhejiang, Shandong, and Guangdong have higher inputs and higher outputs, while the majority of provinces have lower inputs and correspondingly lower outputs for the biomedical industry. It also proves that the development of the biomedical industry requires a higher enterprise R&D investment. The SLR model for operating revenue shows that government investment and enterprise investment have a considerable influence on the development of the biomedical industry. Moreover, enterprise R&D investment is the most important influential factor. The results of the spatial feedback effect and the spatial spillover effect are obvious, indicating that there is a close relationship between biomedical industry development and the degree of investment and development in the neighboring provinces.

Acknowledgement

Funded by: Zhejiang Province, China (2022C03124).

Table 4. SLR model results for operating revenue

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std.Error</th>
<th>t-Statistic</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>W_Revenue</td>
<td>0.257332</td>
<td>0.0992286</td>
<td>2.5933</td>
<td>0.00951</td>
</tr>
<tr>
<td>CONSTAN</td>
<td>24.736</td>
<td>69.2168</td>
<td>0.3573</td>
<td>0.72081</td>
</tr>
<tr>
<td>Government</td>
<td>0.00018029</td>
<td>0.00013962</td>
<td>12.912</td>
<td>0.00000</td>
</tr>
<tr>
<td>Enterprise</td>
<td>0.0213914</td>
<td>0.00625523</td>
<td>3.4197</td>
<td>0.00063</td>
</tr>
</tbody>
</table>

$R^2=0.945368, \log \text{likelihood}= -223.359, \text{Akaike info criterion (AIC)}=454.718, \text{Schwarz criterion (SC)}=460.824, \text{Likelihood Ratio Test}=5.8032(p=0.016)$

References