Academic Journal of Humanities & Social Sciences
ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

Research on regional economic vitality based on ordered logistic regression model

Hong Guo, Xinyue Zhang

Institute of Economics Hebei University, Baoding City Heibei Province 071000

ABSTRACT. *On the road to becoming an economic power, we must pay more attention to the regional and regional individualized development. We should actively promote regional coordinated development, stimulate regional economic vitality and optimize the spatial distribution of modern economy. In recent years, a series of major decisions on the coordinated development of western China, the revitalization of northeast China, the rise of central China and the first development of eastern China have continued. With the implementation, China's road of regional coordinated development has become wider and wider. There is a need to develop a model to quantitative analysis the factors and measures of regional economic activity. Firstly, we design a model to measure the regional economic vitality, and sets up a comprehensive index system of the first and the second level, using Topsis comprehensive evaluation method to analyze and evaluate ten provinces (cities) in the eastern region from different dimensions. Then, an ordered logistic regression model is established. We have taken into account the impact of demographic and entrepreneurial dynamism trends on regional economic dynamism. A total of 6 indicators are selected for model testing and analysis at the population and enterprise levels. The conclusion is that the higher the score, the greater the regional economic vitality. Finally, we have conducted a comprehensive test of the models involved in this article to achieve a better measure of regional economic activity.*

KEYWORDS: *Regional economic vitality, Logistic regression model, Index system, Measure activity*

1. Overview

1.1 Restatement

The economic vitality of the region (province, city) is an important component of the overall competitiveness of the region. In recent years, in order to increase economic vitality, some regions have introduced many preferential policies, such as reducing investment approval procedures, providing financial support for start-ups and lowering the settlement threshold. However, due to the different distribution of urban resources, these policies play different roles in different regions. How to grasp

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

the key factors to improve the economic vitality of the city is worth studying. In order to study how to increase the dynamism of the regional economy, the following problems were solved through surveys and data collection:

(1)The regional economic vitality is influenced by many factors. Taking the eastern region as an example, this paper establishes a reasonable model of influencing factors of economic vitality, draws up a plan to improve the regional economic vitality, and analyzes the influence of the change of regional economic vitality from two angles of population trend and enterprise vitality.

(2)According to the areas discussed in the model, development proposals are put forward to enhance regional competitiveness and maintain sustainable economic development in the region.

1.2 Assumptions

In order to make our model more reasonable, we made the following assumptions:

(1) Use the entropy weight method to calculate the weight of each indicator. This weighting method reflects the objectivity of data processing before using Topsis.

(2) We assume that there is no multicollinearity among the independent variables, and we test it later.

(3) In the search and selection of data, we plan to use continuous variables.

(4) We assume that no major natural disasters will occur in the selected area during the estimation period

2. Analysis of the impact of changes in population dynamics and corporate vitality on regional economic vitality.

2.1 TOPSIS Comprehensive Evaluation Method

In the process of measuring the degree of regional economic activity, we selected multiple indicators for TOPSIS comprehensive evaluation, and finally obtained the closeness of the economic activity of several different cities in the eastern region in the TOPSIS analysis, and analyzed the economic activity of different regions. The degree of economic vitality of the ten cities is sorted according to the closeness value, and the degree of activity is finally divided.

2.1.1 Index selection

First, we need to select indicators to assess the economic vitality of the region, taking into account several key aspects such as population, education, policy, innovation, business environment and economy.

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

In terms of population, we use the indicator of the number of years of home ownership per capita. The number of years of purchase is related to local house prices, which indirectly reflects the extent of economic activity in the region. It measures how many years a person can work in an area before buying a home. This indicator has important demographic relevance and is related to the ability of young people to buy a home and the main survival guarantee for housing problems. Real estate prices are also an important consideration in starting a business. Prices are too high, rents are too high and many companies may abandon strategic plans to set up branches in cities.

$$\text{average number of years of home ownership per capita} = \frac{\text{per capita living space} \times \text{average local price}}{\text{per capital annual income}}$$

In the field of education, we have selected three representative indicators. The average number of years of schooling per capita is the average of the total number of years of schooling for a particular age group. The teacher-student ratio is the teacher-student ratio of high school. The proportion of the population with higher education is the proportion of the population with education as a percentage of the total population in the region.

In terms of policy, the indicator we have chosen is the average rate of growth of revenue from the general public budget.

Representative indicators of innovation include investment in research and development.

Business Environment we choose the road logistics carrying capacity and traffic passenger flow. The disposal rate of domestic waste is the proportion of the treated domestic waste to the total quantity of domestic waste. The sewage treatment rate is the percentage of the total amount of sewage discharged into the Sewage Treatment through the network.

GDP per capita is a dynamic index to reflect the level of economic development in a certain period of time, and is also a basic index whether a country's economy is a dynamic index.

2.1.2 Analysis Process

In the end, we selected 12 indicators to conduct a TOPSIS comprehensive analysis of the economic activity of different cities in the eastern region. The indicator system diagram is as follows:

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

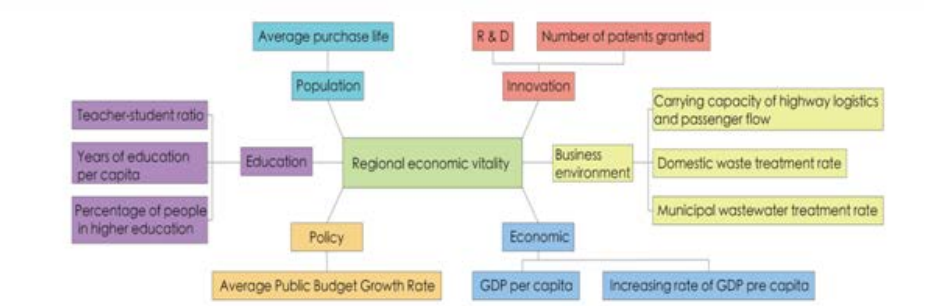

Figure 1: indicator system diagram

Because different indicators have different degrees of response to the economic activity of the region, in order to better evaluate, we use the entropy weight method to calculate the weight of the 12 indicators in the evaluation.

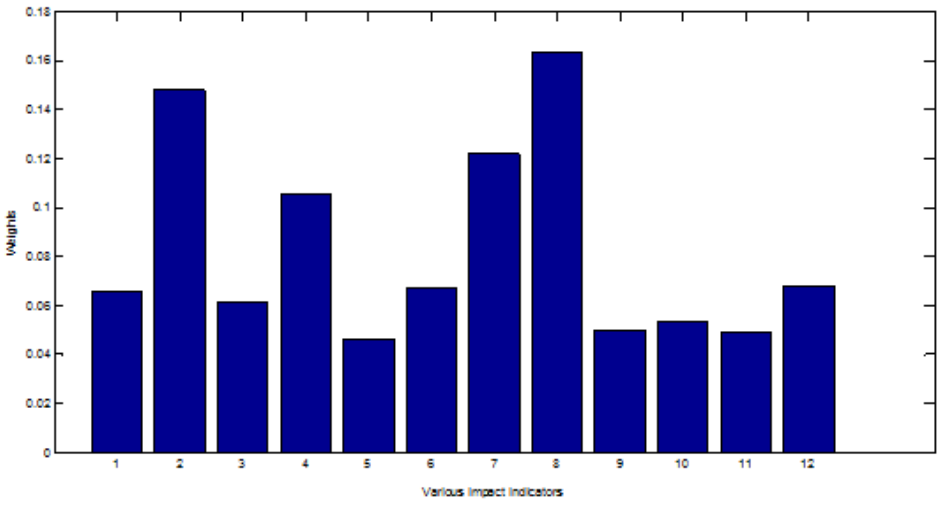

Figure 2: weights of different indicators

After obtaining the weights, we establish a TOPSIS comprehensive evaluation model.

Input parameters:

Weights 0.0661 0.1479 0.0615 0.1056 0.0465 0.0674 0.1218 0.0495 0.0536 0.0491 0.0682

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

x is a matrix of 10 rows and 12 columns. Rows represent cities in the eastern region. Columns are selected indicators.

Output result:

0.1459 0.0443 0.0561 0.1181 0.1112 0.6789 0.2846 0.1121 0.4320 0.0024

They are a comprehensive indicator of the economic activity of Beijing, Hebei, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, and Hainan.

Suppose 12 indicators are expressed as I = I (i), weights are expressed as W = W (i), i = 1, 2, 3 ··· 12, and the degree of urban activity is expressed as T = T (i, j), representing jth The i-th index of each city, j = 1, 2, 3 ··· 10.

$$T\ (i,j) = \sum_{i=1}^{12} W(i)I(i)$$

T is the last value.

2.1.3 Output results

The output histogram is as follows, (regions are arranged from 1 to backward)

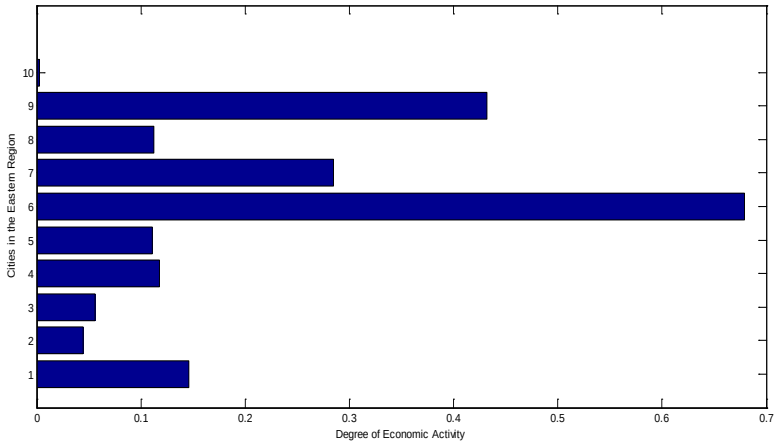

Figure 3:the output histogram

Import the data into the excel table for sorting, you will get a sort of city economic activity, and finally divide the economic activity level as

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

Table 1: the economic activity level

T Value	the Numbers of Cities	Degree of Economic Activity
0-0.1	3	Inactive
0.1-0.2	4	Not very active
0.2-0.3	1	General active
More than 0.3	2	Very active

3. Ordered Logistic Regression

3.1 Problem statement

Researchers in the team want to investigate the degree of people's approval of "regional economic vitality": inactive, expressed by "3"; less active, expressed by "2"; active, expressed by "1"; strongly active,expressed by "0".

In addition, the researchers also investigated some other situations, including: employment rate, dependency ratio, local per capita housing capacity, number of surviving enterprises, fixed investment, and financing scale corporate bonds.

3.2 Overview of the ordered logistic regression model

When using ordered logistic for regression analysis, 4 assumptions need to be considered.

• Hypothesis 1: The dependent variable is unique and it is an orderly multi-categorical variable. For example, the comprehensive competitiveness level of a city can be divided into high, medium, and low; the life satisfaction of a city resident can be divided into satisfactory, general satisfaction, dissatisfaction, etc. The treatment effect of the disease is divided into cured, effective and ineffective.

• Hypothesis 2: There are one or more independent variables, which can be continuous, ordered multi-class, or unordered categorical variables.

• Assumption 3: No multicollinearity between independent variables.

• Assumption 4: The model meets the "proportional advantage" assumption. This means that no matter where the splitting point of the dependent variable is, the influence of each independent variable on the dependent variable in the model does not change, that is, the regression coefficient of the independent variable on the dependent variable has nothing to do with the splitting point.

The principle of ordered and multi-class Logistic regression is to divide multiple categories of the dependent variable into multiple binary logistic regressions in turn. For example, in this article, there are 4 levels in the division of the "regional economic activity" of the dependent variable. Split into three binary logistic

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

regressions: (0 vs 1 + 2 + 3), (0 + 1 vs 2 + 3), and (0 + 1 + 2 vs 3), all of which are lower and higher Compared.

In ordered multi-class logistic regression, it is assumed that in several binary logistic regression, the coefficients of independent variables are equal, only the constant term is different, and the result only outputs the coefficients of a group of independent variables. Therefore, ordered multiple Logistic regression models must test the assumption of equality of independent variable coefficients (the "proportional advantage" hypothesis)(also known as the parallel line test) . If this assumption is not satisfied, consider using unordered multi-class logistic regression.

3.3 Judgment on Hypothesis

Hypothesis 1-2 is a hypothesis of research design, and researchers need to judge based on the research design, so the data hypothesis 3-4 is mainly tested here.

(1)Test hypothesis 3: no multicollinearity between independent variables

Coefficientsa

Model		Collinearity Statistics	
		Tolerance	VIF
1	employment rate	.128	7.785
	Dependency ratio	.108	9.292
	Local per capita purchasing power	.114	8.777
	Number of surviving companies	.220	4.546
	Fixed investment	.118	8.510
	Financing scale corporate bonds	.278	3.595

a. Dependent Variable: Economic vitality

In this example, the tolerances are greater than 0.1, and the variance expansion factors are less than 10, so there is no multicollinearity.

(2)Test result of hypothesis 4

Before explaining the results, we need to look at the test results of Hypothesis 4 (the results of the parallel line test).

The test of parallel lines in the model used in this chapter is $\chi 2 = 0$, and $P = 1$, which shows that the assumption of parallelism holds, that is, the regression equations are parallel to each other and can be analyzed using an ordered logistic process.

Test of Parallel Linesb

Model	-2 Log Likehood	Chi-Square	df	Sig.
Null Hypothesis	.000			
General	.000a	.000	12	1.000

The null hypothesis states that the location parameters(slope coefficients)are the same Across response categories.
a.Link function:Logit.

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

P = 1.000, indicating that the hypothesis of parallelism holds, that is, the regression equations are parallel to each other and can be analyzed using an ordered logistic process.

(3)Goodness of fit test results

The following figure shows the results of the goodness-of-fit test, which are the two goodness-of-fit tests of Pearson and Deviance, respectively. In this example, the result of the Pearson test is $\chi 2 = 0.029$, the deviation = 0.059, and P = 1> 0.05, indicating that the Deviance test result is good for the model pseudo-effect.

Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	.029	21	1.000
Deviance	.059	21	1.000

Link function:Logit。

(4)Model fitting information

The result of Model Fitting Information is a likelihood ratio test on whether the partial regression coefficients of all independent variables in the model are all zero. The result was $\chi 2 = 25.597$ (the difference between the -2 Log Likelihood value of the model with only constant terms and the final model).

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept	25.597			
Final	.000	25.597	6	.000

Link function:Logit。

3.4 In conclusion

Ln (economic vitality oddscat. ≤ inactive) = -924.436 + 972.202 * employment rate -118.148 * dependency ratio + 0.584 * local per capita housing capacity-0.187 * number of surviving companies + 0.01 * fixed investment-0.01 * financing scale-corporate bonds

Ln (economic vitality oddscat. ≤ not very active) = -910.214 + 972.202 * employment rate -118.148 * dependency ratio + 0.584 * local per capita purchase capacity -0.187 * number of surviving companies + 0.01 * fixed investment -0.01 * financing scale-corporate bonds

Ln (economic vitality oddscat. ≤ general active) = -899.738 + 972.202 * employment rate -118.148 * dependency ratio + 0.584 * local per capita house purchasing capacity-0.187 * number of surviving enterprises + 0.01 * fixed investment-0.01 * financing scale-corporate bonds

Through the above analysis, it is not possible to directly give the OR value and its 95% CI. The OR value is obtained by using the Generalized Linear Models

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

module. The orderly Logistic regression is used to analyze the regional economic vitality and employment rate, dependency ratio, and local per capita housing purchase. Utility relationship between capacity, number of surviving companies, fixed investment, financing scale corporate bonds.

The results of the parallel line test are $\chi 2 = 0$, $P = 1$, indicating that the proportional advantage hypothesis exists. Deviance goodness-of-fit test showed that the model fits well, $\chi 2 = 0.59$, $P = 1$, but there are most (75.0%) cells with frequency 0. Model fitting information shows that this model is superior to models with only constant terms.

From the perspective of the number of surviving enterprises, it is considered that it leads to economic vitality is 1.840 times (95% CI: 1.327E139-2.553E139) the value of the inactive economic vitality, $\chi 2 = 0$, $P = 0.997$. In the control group with very active economic vitality, it is considered that the OR value of fixed investment's effect on economic vitality is 0.997 times (95% CI: 10.292-0.097), $\chi 2 = 0$, $P = 0.998$; financing scale corporate bonds are considered The degree of impact on economic vitality is 1.001 times the general active OR value (95% CI: 382.745-0.003), $\chi 2 = 0$, $P = 1$.

4. Sensitivity analysis of the model

For the sensitivity analysis of the model, we choose a one-factor sensitivity analysis to explore the changes in regional economic vitality by changing the per capita GDP growth rate.

The growth rate of per capita GDP increased by 1%, and the changes in the regional economic vitality were between plus and minus 0.3%

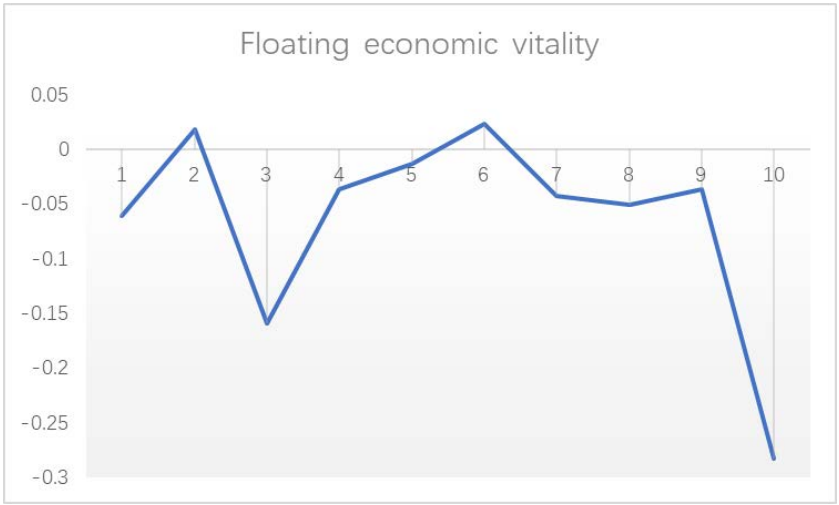

Figure 4:changes in regional economic vitality with a rise of 1%

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

The per capita GDP growth rate is reduced by 1%, and the changes in the regional economic vitality are between plus and minus 0.1%

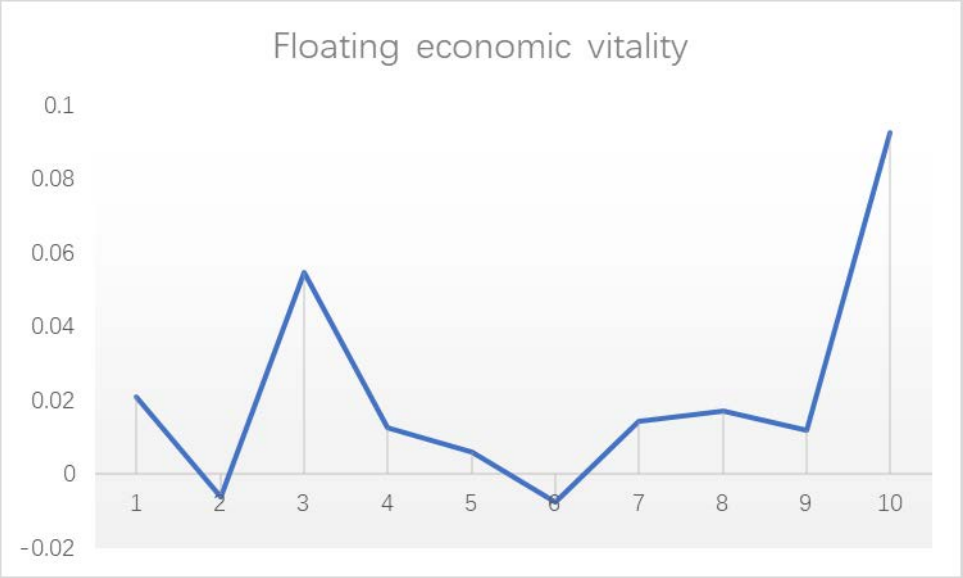

Figure 5:changes in regional economic vitality with a decline of 1%

This shows that our model is reasonable.

5. Development proposals

In this article, we carefully consider the concept of regional economic vitality definition, characteristics and impact factors. We selected representative indicators from seven dimensions of population, business, education, the economy, policy, innovation and business environment, and developed an analytical model. In this paper, a comprehensive index system is established, which involves all aspects of regional economic vitality. Through the logistics model, we can measure the regional economic activity accurately based on the logistic regression equation, the score, the comprehensive score and the calculation of the regional activity ranking table. The regional economic vitality can be estimated by the index data of different regions, and different local governments should develop regional economy according to local conditions, make up the shortage and combine the characteristics of regional economy.

ISSN 2616-5783 Vol.3, Issue 3: 41-51, DOI: 10.25236/AJHSS.2020.030306

References

[1] Lim T. Population Aging in Korea:Macroeconomic Impacts and Financing National Health Insurance[J]. Korean Economic Review, 2016, 32(2):355-382.
[2] Ryan Decker, John Haltiwanger, Ron Jarmin and Javier Miranda, “Where Has all the Skewness Gone? The Decline in High-Growth(Young) Firms in the U.S.”, European Economic Review, Vol.86:4-23, 2016.
[3] Greenlaw.D.A.Kashyap,and K.Schoenholtz. 2012. Stressed Out:Macroprudential Principles for Stress Testing, Chicago Booth Paper No.12-08.