An empirical study on the factors affecting China's fiscal revenue based on regression model

Kun Lin

Audit Bureau of Jinniu District, Chengdu, Sichuan610000, China

Keywords: financial revenue, DW test, Goldfeld-Quanadt test, White test

Abstract: There are various types of factors affecting a country's fiscal revenue, such as social electricity consumption, tax revenue, CPI output value of the three major industries, PPI, fixed asset investment, employment, and so on. Various related factors influencing China's fiscal revenue, the author of this paper constructs the econometric model based on regression algorithm, using Wind database, gathering information using Eviews8.0 software to establish regression model, data and autocorrelation, heteroscedasticity and multicollinearity analysis, analyzes the main factors affecting China's fiscal revenue and the influence of various factors. Combined with the knowledge of economics, the paper gives a reasonable explanation to the model. Finally, some constructive policy Suggestions are put forward.

1. Introduction

Fiscal revenue for the country or region with special powers of the government, government by law, in a certain period of time (like years) to obtain the sum of the various forms of income, including foreign exchange income, business income and debt income, income of tax revenue, key construction funds, and state funds equity income, fees income, debt, income, income of fines, etc. The level of a country's fiscal revenue is an important indicator of a country's economic strength. In a certain period of time, the size of fiscal revenue is affected by many factors, such as GDP, GNP, employment of the whole society, foreign exchange income scale, tax scale and fine scale.

Fiscal revenue is important revenue of a country, which represents the hard power of a country. With high fiscal revenue, the country will have more money to do more things. We can invest in infrastructure and build railways, public transport, etc. This will improve people's livelihood and raise their living standards and satisfaction with life. The principal contradiction facing society today is that the growing material and cultural needs of the people are in conflict with the imbalanced productive forces and inadequate material and cultural life. The growth of fiscal revenue can partially solve the above problems. This is very important.

Fiscal revenue is one of the main means of national macro-control. It can optimize the allocation of various resources, thus stimulating regional economic development and improving the average living standard of the people. In the past three decades, with the reform of the economic system and the rapid development of the economy, China's fiscal revenue has undergone many changes. In order to deeply explore the important factors affecting the growth of fiscal revenue, analyze the essential laws behind the increase of fiscal revenue, and predict the future growth characteristics of China's fiscal revenue, it is necessary to build an appropriate econometrics algorithm model.

There are many factors affecting fiscal revenue, as many as hundreds, or even thousands. Data mining algorithms such as machine learning and deep learning can be used to explore the influencing factors of financial revenue and the development law behind them. However, machine learning algorithms, especially deep learning algorithms, require a large amount of data, which is not suitable for a small amount of data. Moreover, the interpretability of machine learning algorithm is not strong. At the same time, the econometrics algorithm has a strong explanatory ability for practical problems, which can effectively explain and analyze the internal relations of various factors affecting financial revenue and the importance of their respective influences.

Finance is the behavior of the government and the concrete embodiment of its functions. It mainly has three functions: resource allocation, income redistribution and macro-economic control. Fiscal revenue is the public revenue of our government, which is the capital need in the distribution of national income to ensure the government to exercise government functions carry out public policies and export external services. The situation and increasing trend of fiscal revenue are related to the progress of society and the development of national economy. Therefore, it is very necessary to study the growth of financial revenue and its influencing factors and the law of development. Then, based on the fiscal revenue for the dependent variable (y), with tax revenues (x2), number of employees (x3), fixed asset investment (x4), a total of four indicators as independent variables, using Eviews8.0 multivariate linear regression analysis, autocorrelation, heteroscedasticity and multicollinearity test, the factors that affect the revenue model was constructed, deep analyze the main factors affecting China's fiscal revenue, for the healthy development of our country finance income situation put forward constructive Suggestions and measures.

2. Research status and background

2.1 Research Status

There are not many researches on fiscal revenue from domestic and foreign scholars. Mostly at a general level. From the perspective of literature research and demonstration, there are many researches. However, there are few researches on the model level. Many scholars have studied the contribution of taxes and fiscal revenue to the country. This paper discusses the function, influence and public function of fiscal revenue.

Some scholars have made model analysis on the factors affecting fiscal revenue, but the model fitting and prediction are not accurate. In addition, there may be autocorrelation, heteroscedasticity and multicollinearity, resulting in poor model analysis and interpretation! The prediction of the model is not good either!

Some scholars tried to use machine learning algorithm to study the influencing factors of financial revenue, but failed. Because the amount of data is too small, machine learning does not have enough data to find the statistical laws of each feature, so it is unable to use each feature to model with machine learning algorithm. Because machine learning algorithms require a certain amount of data. Too little data won't do. Deep learning algorithms, such as artificial neural network ANN, require a larger amount of data. Unless the amount of data affecting our fiscal revenue is large enough. China was founded only in 1949, and until 2020, there are at most 71 years of data. So that's 71 rows. The amount of data is really small.

In 2019, some scholars tried to study the influencing factors of fiscal revenue from the perspective of correlation analysis. Fiscal revenue was set as the dependent variable (Y), and other variables, such as GDP, GNP, tax revenue, number of employed persons, sea freight index, CPI, PPI, and social electricity consumption, were set as independent variables for cross-correlation analysis. Investigate which of the above factors have a high correlation with fiscal revenue (Y), so as to predict the factors that have a strong correlation with fiscal revenue. However, no relevant model has been established.

Some scholars have written relevant summary papers, specifically discussing relevant knowledge theories of fiscal revenue, demonstrating and analyzing from the perspective of literature review, and analyzing the function, responsibility, role and influencing factors of fiscal revenue. Analyze the development status and history of financial revenue, and conduct in-depth analysis and research.

Some scholars have made a knowledge mapping expert system on the relationship between financial revenue and other factors. To sort out the direct or indirect relationship between financial revenue and other influencing factors, as well as the horizontal and vertical relationship.

Some scholars, based on the time series data of China's annual fiscal revenue, have made an autoregression model based on ARMA and ARIMA to fit and forecast the time series data of fiscal revenue. This is a time series analysis of a single factor, not a multi-factor analysis of influencing factors.

Based on the data of China's fiscal revenue and GDP, some scholars constructed a model based on binary linear and nonlinear regression based on the above two data, with fiscal revenue as Y and

GDP as X. The final fitting effect was not bad. It plays a good role in predicting the future development trend of China's fiscal revenue.

Some scholars, from the macro, economic and policy perspective, based on empirical analysis and economic development cycle theory, predict the future development trend of China's fiscal revenue. However, the prediction effect is not very good. After all, without models and algorithms, the historical data cannot be well fitted to predict the future data.

2.2 Background

Since the reform and opening up, China's economy has made considerable progress, and the national economy has entered the fast lane. These benefits from the opportunity of the party's policy and reform and opening up. China has made brilliant achievements in politics, economy, finance, culture and sports. China has gradually grown from the 10th largest economy in the world to the 4th largest economy, the 3rd largest economy, and now the 2nd largest economy. By the end of 2019, China's economic aggregate had reached more than 60 percent of that of the United States.

The living standards of the people have been gradually improved, with a prosperous life and a prosperous economy. The Chinese people have also stepped into a well-off society. The national economy has developed by leaps and bounds. Infrastructure development, medical reform and welfare, public medical care, and various social security systems and measures have also been gradually implemented, enabling the people to enjoy benefits in person. All of the above cannot be separated from the guarantee of China's fiscal revenue, with which the country can concentrate its efforts to do great things and better promote the development of China's economy and the improvement of people's living standards. Therefore, it is of great significance to study the financial revenue and its influencing factors in depth.

Theoretical significance: National fiscal revenue is of great significance, and all government work cannot operate without effective financial support. A nation is not a nation without revenue.

First, fiscal revenue is the financial guarantee for the realization of state functions. Without fiscal revenue, the state has no support and can do nothing. Unable to concentrate on doing great things, unable to better optimize the allocation of resources, do more things; Various social security measures and social welfare will be out of the question. There will be no substantial improvement in people's living standards. Fiscal revenue is the material premise, with fiscal revenue, the functions of the state can be smoothly performed. The implementation of policies, regulations and public functions at all levels can be better implemented.

Second, fiscal revenue is an important way to correctly handle and balance material interests. There are many ways in which a country USES its fiscal revenue, which is supported by politics, economy, infrastructure, science and technology, culture, sports, public health and medical care. How to allocate and optimize resources is a very important content. In order to make better use of fiscal revenue and maximize the use of fiscal revenue, it is very important to deal with and measure the relationship of material benefits. Only the rational allocation of resources can give full play to the important role of fiscal revenue. In order to better and faster promote the development of the national economy and the improvement of people's living standards.

Third, fiscal revenue is the premise of fiscal expenditure. Had financial revenue, just have money on paper, this is the premise that finance spends safeguard. Fiscal revenue refers to the collection of national tax revenue, penalty revenue, and national infrastructure fund revenue into the national Treasury, which is collected together to form fiscal revenue. Thus, for the fiscal expenditure plan formed the fund guarantee. Fiscal expenditure plan is distributed according to fiscal revenue, so. Fiscal revenue is the premise of fiscal expenditure. Without fiscal revenue, fiscal expenditure becomes a ghost without a source of water or a tree without roots.

Fiscal revenue is the material guarantee and prerequisite of a country's fiscal expenditure. A country's fiscal revenue is an important reflection of its economic strength. A country with a large fiscal revenue proves that it has strong economic strength and can concentrate on doing great things to better build the whole country. Fiscal revenue is also an important financial guarantee for the state to carry out macro-control and bring economic leverage into full play. To carry out macro-control, the state needs to allocate a certain amount of capital and bring into full play the role of leverage. Fiscal income distribution is one of the main ways to adjust the framework of primary distribution of national income and realize fair distribution of national wealth. To a certain extent, it is beneficial to reduce the gap between the rich and the poor, reduce the imbalance of the first distribution of social wealth, and facilitate the second rational distribution of wealth. Conducive to the smooth operation of the national economy.

The reasonable distribution of fiscal revenue is an important lever of national economy. It is one of the important ways to adjust the balance of social aggregate demand and supply. It is one of the important means for the state to participate in economic adjustment and an important guarantee for national economy and people's livelihood and national infrastructure construction.

Therefore, it is very important to deeply analyze the financial revenue and the important factors affecting the financial revenue. We can carry out the analysis from the two dimensions of qualitative and quantitative analysis. With the help of econometric models and relevant testing means, such as T test, F test, autocorrelation, heteroskedasticity, multicollinearity test, DW test, Goldfeld-Quanadt test and White test, it will be helpful for us to carry out quantitative analysis more scientifically. It is more helpful for us to analyze the factors that affect our fiscal revenue very thoroughly and clearly.

3. The quantitative analysis of financial revenue influencing factors

3.1 Variable Selection

There are many factors that can affect fiscal revenue, and complicated, such as GDP, GNP, CPI, PPI, strength index, social power consumption, tax revenues, social workers, books, national assets total amount of investment, word of shipping, real estate investment, real estate prices, the Shanghai index, the hang seng index, rebar prices, etc., there are about 200.However, due to the limited space, this paper can only select some of the above factors to investigate the extent of their impact on financial revenue and the applicable model. After screening, this paper selects four factors of GDP, tax revenue, number of social workers and fixed asset investment for analysis. Historical data is obtained from the Wind database.

3.2 Data Description

(1) Revenue: is the normal operation of the machine, material support, the government in order to maintain its normal operation and implement national public functions to the enterprises and institutions, factories, institutions and organizations such as collection of taxes, tax and national infrastructure construction fund, the sum of income of fine, is an important material guarantee fiscal revenue and expenditure. t is an important indicator and symbol to measure a country's economic development strength. It is one of the important tools for the state to participate in macroeconomic regulation and control. It can bring into play the power of the state's economic lever to better carry out regulation and control measures and serve the real economy. Is an important wealth of a country, but also participate in the second distribution of the country, to achieve fair distribution between different groups; It is one of the important tools to reduce the gap between rich and poor.

(2) GROSS domestic product (GDP) : it is an important indicator to measure the economic strength of a country and represents the comprehensive national strength of a country; It is the sum of all the output value of national production produced by all enterprises, institutions and self-employed households in a country within a certain period of time. It reflects a country's comprehensive national strength and economic development level, is an important development index.

(3) Investment in fixed assets: refers to the economic activities of building and purchasing fixed assets, that is, investment and construction activities of fixed assets; It reflects the strength and scale of a country's investment in fixed assets, and measures the importance a country attaches to investment in fixed assets and the proportion of investment in fixed assets.

(4) Number of employees: the number of economically active people in the population. It does not

include retired, disabled and domestic workers, or other persons unable to participate in activities, etc.It reflects the overall size of the population engaged in economic production in a country.

(5) Tax revenue: Tax revenue is a kind of revenue in kind or money collected by the state from enterprises, public institutions, self-employed households and other entities within the territory for free in accordance with the provisions of the tax law and certain standards.It includes a variety of categories, including business tax, value-added tax, value-added tax, individual income tax, and so on.It is the adjustment of high income, is the important guarantee of the national secondary distribution.

4. Model Establishment

4.1 Model Description

There are many factors affecting fiscal revenue, including the following aspects: tax revenue, profit revenue turned over by state-owned enterprises, debt revenue, expenses and other income, among which tax revenue is an important part of fiscal revenue.On the other hand, fiscal revenue is also affected by GDP, the average number of employees, the total fixed asset investment in the country and many other factors. Here, GDP changes can be used to illustrate the impact of changes in other factors other than taxation on fiscal revenue.

4.2 Model data description

The data in this research report are collected from the statistical Yearbook of The National Bureau of Statistics of China from 1987 to 2019

Table 1: The national fiscal revenue and related data from 1987 to 2019 are attached

The index name	National revenue	GDP: the present price	National public revenue	The annual average number of all employees: the national total	The total fixed asset investment of the whole society
frequency	years	years	years	years	years
unit	One hundred million yuan	One hundred million yuan	One hundred million yuan	Ten thousand people	One hundred million yuan
source	National Bureau of Statistics	National Bureau of Statistics	The Ministry of Finance	National Bureau of Statistics	National Bureau of Statistics
1987	2199.40	12058.60	2140.36	52783.00	3791.70
1988	2357.20	15042.80	2390.47	54334.00	4753.80
1989	2664.90	16992.30	2727.40	55329.00	4410.40
1990	2937.10	18667.80	2821.86	64749.00	4517.00
1991	3149.48	21781.50	2990.17	65491.00	5594.50
1992	3483.37	26923.50	3296.91	66152.00	8080.10
1993	4348.95	35333.90	4255.30	66808.00	13072.30
1994	5218.10	48197.90	5126.88	67455.00	17042.10
1995	6242.20	61339.90	6038.04	8360.00	20019.30
1996	7407.99	71813.60	6909.82	8187.00	22913.50
1997	8651.14	79715.00	8234.04	7873.00	24941.10
1998	9875.95	85195.50	9262.80	6195.81	28406.20
1999	In 11444	90564.40	10682.58	5805.05	$29854.70
2000	13395.23	100280.10	12581.51	5559	32917.73
2001	16386.04	110863.10	15301.38	5441.43	37213.49

2002	18903.64	121717.40	17636.45	5520.66	43499.91
2003	21715.25	137422.00	$20017.31...	5748.57	55566.61
2004	26396.47	161840.20	24165.68	6622 15	70477.40
2005	31649.29	187318.90	28778.51	6895.96	88773.60
2006	38760.20	219438.50	34804.35	7358.43	109998.20
2007	51321.78	270092.30	45621.97	7875.20	137323.94
2008	61330.35	319244.60	54223.79	8837.63	172828.40
2009	68518.30	$348517.70	59521.59	8831.22	224598.80
2010	83101.51	412119.30	73210.79	End at 9544.71	251683 first
2011	103874.43	487940.20	End at 89738.39	9167.29	311485.13
2012	117253.52	538580.00	100614.28	9567.32	374694.74
2013	129209.64	592963.20	$110530.70	9791.46	446294 15
2014	140370.03	643563.10	$119175.31...	Of 9977	512020 index
2015	152269.23	688858.20	124922.20	9775.02.	561999.83
2016	159604.97	746395.10	130360.73	9475.57	606465.66
2017	172592 first	832035.90	144369.87	8957.89	End at 641238.39
2018	183351.84	919281.10	156402.86	7942.30	645675.00
2019	190382.23	990865.10	157992.20	8233.19	560874.00

4.3 Model establishment

The linear regression model was established with the national fiscal final revenue as the explained variable and GDP (current price), various taxes in the national fiscal final revenue, number of employees at the end of the year, and the total fixed asset investment of the whole society as the explanatory variables:

$$Y_t = beta._0 + beta._1 X_{1t} + beta._2 X_{2t} + beta._3 X_{3t} + Beta._4 X_{4t} + u_i$$

Among them, Y_t -- National revenue X_{1t} - GDP: the present price

X_{2t}-- National Public Revenue X_{3t}-- The annual average number of all employees: the national total

X_{4t} -- The total amount of fixed asset investment completed by the whole society

$Beta._0$, $beta,_1$, $beta,_2$, $beta,_3$, $beta,_4$, $beta,_5$-- Represents the undetermined coefficient

u_i -- Represents the random error term

4.4 Regression model

The following results can be obtained by using Eviews software and OLS regression method

Table 2: OLS regression results

Dependent Variable: Y				
Method: Least Squares				
Date: 06/20/20 Time: 07:59				
Sample: 1987 2019				
Included observations: 33				
Variable	Coefficient		t-Statistic	Prob.
C	1368.763	0.008282	1.878368	0.0708
X1	0.020893	0.056463	2.522665	0.0176
X2	1.191468	0.013947	21.10174	0.0000
X3	0.023822	0.008544	1.708051	0.0987
X4	0.021827		2.554790	0.0163
R-squared	0.999863	S.D. dependent var		22220.32
Adjusted R-squared	0.999843	Akaike info criterion		31238.58

S.E. of regression	391.5801	Schwarz criterion	14.91698
Sum squared resid	4293379.	Hannan-Quinn criter.	15.14373
Log likelihood	241.1302	Durbin-Watson stat	14.99328
F-statistic	50906.31		1.799578
Prob(F-statistic)	0.000000		

Y = 1368.763 0.020893 X1 + 1.191468 X2 X3 0.023822 + 0.021827 X4

T = (1.878368) (-2.522665) (21.10174) (-1.708051) (2.554790)

R^2 DW = 50906.31 0.999863 = = = 1.799578 $\overline{R}^2$ 0.999843 F

5. Model Test

5.1 Economic Test

The results of the model estimate show that, assuming that other variables remain unchanged, every 1% increase in GDP (current price) reduces national fiscal revenue by 2.893% on average. Assuming that other variables remain unchanged, tax revenue increases by 1%, and national fiscal revenue increases by 119.1468% on average. Assuming that other variables remain unchanged, if the number of employees increases by 1%, the national fiscal revenue will decrease by 2.3822% on average. Assuming that other variables remain unchanged, the total fixed asset investment of the whole society will increase by 1%, and the national fiscal final revenue will increase by 2.1827% on average. This is consistent with theoretical analysis and empirical judgment.

5.2 Statistical Test

1) Goodness of fit test

R can be obtained from the data in 3.42 =0.999863, the modified determination coefficient =0.999843, which indicates that the model fits the samples well.

2) Significance test of variables (T test)

For H0: j=0 (j=1,2,3,4,5), given significance level =0.05, the critical value T of N-k-1 =28 was obtained by looking up the T-distribution tableα/2（n-k-1）= 2.048.Corresponding t statistic data is available from 3.4 1.87836, 2.522665, 21.10174, 1.708051, 2.554790, of which the x2, x4 t statistic absolute value is greater than 2.048, shall reject the null hypothesis and x1, x3 t statistic absolute value is less than 2.048, the alternative hypothesis should be rejected, which means tax revenue, total investment in fixed assets, the whole society to be explained variable, respectively, national finance income have a significant impact, and gross domestic product (present price),The number of employees has no significant influence on the national fiscal revenue of the explained variable.

3) Significance test of equation (F test)

For H0:j=0 (j= 2,3,4,5), given significance level =0.05, find the critical value F (3, 28) =2.95 in the F distribution table for k-1=3 and n-k-1=28.From 3.4, F=50906.31. Since F=50906.31> F (3,28) =2.95, the null hypothesis H0: j=0 (j= 2,3,4,5) should be rejected, indicating that the regression equation is significant, that is, gross domestic product (current price), tax revenue, number of employees, total fixed asset investment of the whole society and other variables combined have a significant impact on the national fiscal revenue.

6. Multicollinearity test and its modification

6.1 Multicollinearity check

From the coefficient of regression results and t value, we can see that the model may have multicollinearity. Now we calculate the correlation coefficient of explanatory variables. The correlation coefficient matrix of explanatory variables is as follows:

Table 3: Correlation coefficient matrix

variable	X1	X2	X3	X4
X1	1.000000	0.994618	0.713051	0.985822

X2	0.994618	1.000000	0.652140	0.995642
X3	0.713051	0.652140	1.000000	0.618040
X4	0.985822	0.995642	0.618040	1.000000

According to the correlation values, the explanatory variables are highly correlated, and the model has serious multicollinearity.

6.2 Multicollinearity correction

A stepwise regression method is used to test and solve the multicollinearity problem. The unary regression of Y on X1, x2, X3, and X4 is made respectively, and the unary regression estimation results are as follows:

Table 4: OLS regression results

variable	X1	X2	X3	X4
Parameter estimates	0.215766	1.153796	1.729660	0.315383
T statistic	45.04898	242.1384	4.619980	68.18362
R^2	0.984954	0.999472	0.407766	0.993376
$\bar{R}^2$	0.984469	0.999455	0.388662	0.993162

Sort by size of modified determination coefficient: X2,X4,X1,X3Visible, X2The corrected determination coefficient is the largest and should be X2Step by step regression of other variables is added on the basis.

With X2On the basis, other variables are successively added for stepwise regression. So let's add X first4Regression results are as follows:

Table 5: OLS regression results

Dependent Variable: Y				
Method: Least Squares				
Date: 06/20/20 Time: 8:40				
Sample: 1987 2019				
Included observations: 33				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	210.0655	149.2651	1.407332	0.1696
X2	0.981561	0.041235	23.80436	0.0000
X4	0.047430	0.011306	4.195242	0.0002
R-squared	0.999667	Mean dependent var		22220.32
Adjusted R-squared	0.999645	S. d. dependent var		31238.58
S.E. of regression	588.7962	Akaike info criterion		15.68055
Sum squared reside	10400429	Schwarz criterion		15.81659
Log likelihood	255.7290	F - statistic		45022.32
Durbin-Watson stat	1.192772	Prob (F - statistic)		0.000000

$\hat{Y}_t = \text{-}210.0655 + 0.981561X2 + 0.047430X4$

R t (23.80436) (4.195242) ^{2}= 0.9999667

When alpha =5%, X $t_{\alpha/2}(n-k-1) = t_{0.025}(33-2-1) = 2.042$ $_4$The t test of the parameter is significant, so it is not eliminated, and X is added$_1$Return to:

Table 6: OLS regression results

Dependent Variable: Y		
Method: Least Squares		
Date: 06/20/20 Time: 8:46		

Sample: 1987 2019				
Included observations: 33				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	139.6388	118.4828	1.178557	0.2482
X2	1.236054	0.051696	23.91019	0.0000
X4	0.023651	0.008752	2.702301	0.0114
X1	0.031846	0.005412	1.884012	0.0000
R-squared	0.999848	Mean dependent var		22220.32
Adjusted R-squared	0.999832	S. d. dependent var		31238.58
S.E. of regression	404.3182	Akaike info criterion		14.95549
Sum squared reside	4740724.	Schwarz criterion		15.13689
Log likelihood	242.7657	Kerri Hannan - Quinn criter.		15.01653
F-statistic	63664.69	Durbin - Watson, stat		1.623986
Prob(F-statistic)	0.000000			

$\hat{Y}_t = 139.6388 + 1.236054X2 + 0.023651X4 - 0.031846X1$

T (23.91019) (2.702301) (-1.884012) R^2= 0.999848

When alpha =5%,X $t_{\alpha/2}(n-k-1) = t_{0.025}(33-3-1) = 2.045$ 1The T test of the parameter is not significant, and its coefficient is inconsistent with the economic significance, so it is eliminated and X is added3Return to:

Table 7: OLS regression results

Dependent Variable: Y				
Method: Least Squares				
Date: 06/20/20 Time: 11:02				
Sample: 1987 2019				
Included observations: 33				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2682.357	554.9036	4.833916	0.0000
X2	1.073606	0.034513	31.10762	0.0000
X4	0.025990	0.009125	2.848185	0.0080
X3	0.051063	0.009609	1.314216	0.0000
R-squared	0.999831	Mean dependent var		22220.32
Adjusted R-squared	0.999814	S. d. dependent var		31238.58
S.E. of regression	426.2579	Akaike info criterion		15.06118
Sum squared reside	5269178.	Schwarz criterion		15.24257
Log likelihood	244.5094	Kerri Hannan - Quinn criter.		15.12221
F-statistic	57278.69	Durbin - Watson, stat		1.685456
Prob(F-statistic)	0.000000			

$\hat{Y}_t = 2682.357 + 1.073606X2 + 0.025990X4 - 0.051063X3$

T (31.10762) (2.848185) (-1.314216) R^2= 0.999831

When alpha is equal to 5%, X $t_{\alpha/2}(n-k-1) = t_{0.025}(33-3-1) = 2.045$ 3The t test of parameters is not significant, so it is eliminated.

Finally, the regression results after correcting the severe multicollinearity are as follows:

$\hat{Y}_t$ = -210.0655 + 0.981561$X2$ + 0.047430$X4$

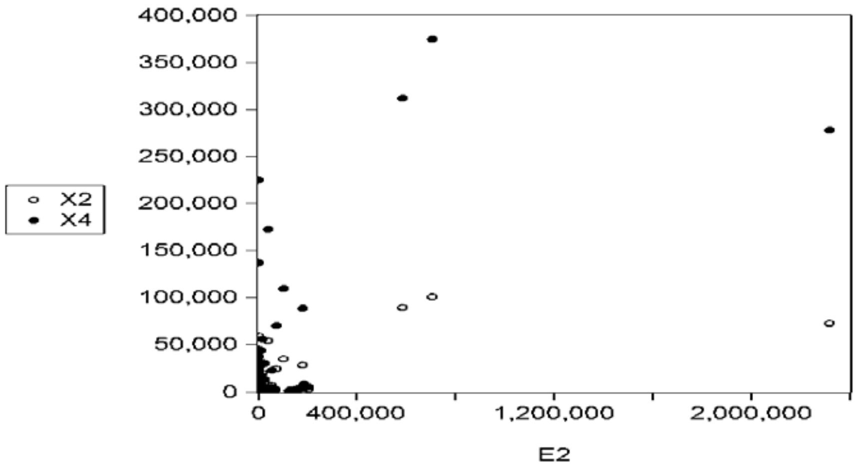
T (23.80436) (4.195242)

R^2DW = 45022.32 0.9999667 = = = 1.192772 $\overline{R}^2$ 0.999645 F

7. Heteroscedasticity test and its modification

7.1 Heteroscedasticity test

7.1.1 Draw the residual squared pair X2, X4The scatter plot



Graph one: x_2 x $_4$A scatter diagram

As can be seen from the figure, the scattered points are concentrated in the lower left corner, and the model may have heteroscedasticity. Next, we use other methods to further verify the existence of heteroscedasticity of the model.

7.1.2 Goldfeld - Quanadt inspection

Since n=33 removes a quarter of the observed values, that is, about 7 observed values, the remaining part is divided equally into two sample intervals: 1987~1999 and 2007~2019, both of which have 13 samples, i.e., N1=n2= 13.OLS was used for estimation.

Regression results of the first set of data (1987-1999) :

Table 8: OLS regression results

Dependent Variable: Y				
Method: Least Squares				
Date: 06/20/20 Time: 11:40				
Sample: 1987 1999				
Included observations: 13				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	746.2393	82.26066	9.071642	0.0000
X2	0.477197	0.109447	4.360068	0.0014
X4	0.145068	0.051156	2.835772	0.0177
R-squared	0.976405	Mean dependent var		2113.554
Adjusted R-squared	0.971686	S. d. dependent var		784.7651
S.E. of regression	132.0501	Akaike info criterion		12.80341
Sum squared reside	174372.3	Schwarz criterion		12.93379
Log likelihood	80.22219	Kerri Hannan - Quinn criter.		12.77662
F-statistic	206.9107	Durbin - Watson, stat		1.281482
Prob(F-statistic)	0.000000			

RSS1=174372.3

Regression results of the second group of data (2007-2019) :

Table 9: OLS regression results

Dependent Variable: Y		
Method: Least Squares		
Date: 06/20/20 Time: 18:17		
Sample: 2007 2019		

Included observations: 13				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2000.494	501.3152	3.990491	0.0032
X2	1.163040	0.060317	19.28214	0.0000
X4	0.004036	0.015298	0.263844	0.7978
R-squared	0.999788	Mean dependent var		53267.57
Adjusted R-squared	0.999741	S. d. dependent var		34089.78
S.E. of regression	548.9453	Akaike info criterion		15.66619
Sum squared reside	2712068.	Schwarz criterion		15.78742
Log likelihood	90.99716	Kerri Hannan - Quinn criter.		15.62131
F-statistic	21206.10	Durbin - Watson, stat		2.019121
Prob(F-statistic)	0.000000			

The sum of squared residuals RSS2=2712068

So F= RSS2/RSS1= 2712068/174372.3=15.5533

Looking up the threshold at a given alpha =5%, $F_{0.05}(9,9) = 3.18$ $F > F_{0.05}(9,9)$

Therefore, the hypothesis that the variance of two groups of subsamples is the same is denied, so that the population random term has increasing heteroscedasticity.

7.1.3 White inspection

Using EViews yields the following results:

Table 10: OLS regression results

Heteroskedasticity Test: White				
F-statistic	5.390677	Prob. F (5, 27)		0.0015
Obs*R-squared	16.48574	Prob. Chi - Square (5)		0.0056
Scaled explained SS	22.69269	Prob. Chi - Square (5)		0.0004
Test Equation:				
Dependent Variable: RESID^2				
Method: Least Squares				
Date: 06/20/20 Time: 18:48				
Sample: 1987 2019				
Included observations: 33				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	358749.1	164260.7	2.184023	0.0378
X_2	191.3157	111.0285	1.723122	0.0963
X_2^ 2	0.013465	0.016216	0.830385	0.4136
X_2*X_4	0.007106	0.008786	0.808835	0.4257
X_4	55.00532	29.51195	1.863832	0.0733
X_4^ 2	0.000936	0.001197	0.781993	0.4410
R-squared	0.499568	Mean dependent var		315164.5
Adjusted R-squared	0.406895	S.D. dependent var		584138.3
S.E. of regression	449864.1	Akaike info criterion		29.03424
Sum squared reside	5.46 e+12	Schwarz criterion		29.30634
Log likelihood	473.0650	Hannan-Quinn criter.		29.12580
F-statistic	5.390677	Durbin-Watson stat		1.902545
Prob(F-statistic)	0.001454			

So you can see nR from this table2 = 16.48574, and under 5% significance level, the critical value X look-up table2(5) =11.07.Because of nR2 = 16.48574 > X2(5) =11.07, therefore, it indicates that the model has heteroscedasticity.

It can be seen from the test results of the above methods that the model has heteroscedasticity, so we will modify it.

7.2 Modification of heteroscedasticity

In the estimation process of weighted least square method, we use w1=1/ SQR (R2),w2=1/ R2 ^2, and W respectively3= 1 / r2, (r2 = resid ^ 2).The weight w was found by comparison3.I'm going to give you the weight w1, w2, w3. Regression results is:

When the weight is W1: the result is as follows

Table 11: OLS regression results

Dependent Variable: Y				
Method: Least Squares				
Date: 06/20/20 Time: 19:04				
Sample: 1987 2019				
Included observations: 33				
Weighting series: W1				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	695.9942	44.68025	15.57722	0.0000
X2	0.492671	0.086759	5.678604	0.0000
X4	0.171294	0.028554	5.998974	0.0000
	Weighted Statistics			
R-squared	0.988723	Mean dependent var		3300.513
Adjusted R-squared	0.987972	S. d. dependent var		788.8441
S.E. of regression	252.2865	Akaike info criterion		13.98552
Sum squared resid	1909455.	Schwarz criterion		14.12156
Log likelihood	227.7610	Kerri Hannan - Quinn criter.		14.03129
F-statistic	1315.194	Durbin - Watson, stat		0.625484
Prob(F-statistic)	0.000000			
	Unweighted Statistics			
R-squared	0.996646	Mean dependent var		22220.32
Adjusted R-squared	0.996422	S. d. dependent var		31238.58
S.E. of regression	1868.534	The Sum squared resid		1.05 e+08
Durbin-Watson stat	0.813248			

WLS regression results were obtained according to the results

$$\hat{Y}_i = 695.9942 + 0.492671X2 + 0.171294X4$$

SE (44.68025) (0.086759) (0.028554)

T (15.57722) (5.678604) (5.998974)

R^2= 0.988723 = 1315.194, p = 0.0000001 $\overline{R}^2 = 0.987972$ F

In order to analyze the correction of heteroscedasticity, after using WLS to estimate each model, the White test should be used to determine whether the model has heteroscedasticity again.

Table 12: White test results

Heteroskedasticity Test: White			
F-statistic	3.825036	Prob. F (5, 27)	0.0095
Obs*R-squared	13.68300	Prob. Chi - Square (5)	0.0178
Scaled explained SS	7.699220	Prob. Chi - Square (5)	0.1736

Take the significance level =0.05, because of nR^2=13.68300>, so we still have the heteroscedasticity. $\chi^2_{0.05} = 11.07$ W1 is not desirable as a weight.

When the weight of is W2: the results are as follows

Table 13: OLS regression results

Dependent Variable: Y				
Method: Least Squares				
Date: 06/20/20 Time: 19:05				
Sample: 1987 2019				
Included observations: 33				
Weighting series: W2				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	747.7831	33.83127	22.10331	0.0000
X2	0.371120	0.127856	2.902630	0.0069
X4	0.212610	0.057214	3.716027	0.0008
	Weighted Statistics			
R-squared	0.930695	Mean dependent var		1505.366
Adjusted R-squared	0.926075	S. d. dependent var		2581.279
S.E. of regression	119.1164	Akaike info criterion		12.48459
Sum squared reside	425661.8	Schwarz criterion		12.62063
Log likelihood	202.9957	Kerri Hannan - Quinn criter.		12.53036
F-statistic	201.4343	Durbin - Watson, stat		1.631015
Prob(F-statistic)	0.000000			
	Unweighted Statistics			
R-squared	0.997055	Mean dependent var		22220.32
Adjusted R-squared	0.996859	S. d. dependent var		31238.58
S.E. of regression	1750.700	The Sum squared resid		91948524
Durbin-Watson stat	1.331100			

WLS regression results were obtained according to the results

$$\hat{Y}_i = 747.7831 + 0.371120X2 + 0.212610X4$$

SE (33.83127) (0.127856) (0.057214)
T (22.10331) (2.902630) (3.716027)

R^2= 0.930695 = 201.4343, p = 0.000000 $\overline{R}^2 = 0.926075$ F

In order to analyze the correction of heteroscedasticity, after using WLS to estimate each model, the White test should be used to determine whether the model has heteroscedasticity again.

Table 14: White test results

Heteroskedasticity Test: White			
F-statistic	6.817626	Prob. F (6, 26)	0.0002
Obs*R-squared	20.17599	Prob. Chi - Square (6)	0.0026
Scaled explained SS	86.51920	Prob. Chi - Square (6)	0.0000

Take the significance level =0.05, because of nR^2Is equal to 20.17599>, so we still have the heteroscedasticity. $\chi^2_{0.05} = 12.59$ W2 for weight is not desirable.

When the weight of is W3: the results are as follows:

Table 15: OLS regression results

Dependent Variable: Y				
Method: Least Squares				
Date: 06/20/20 Time: 19:42				
Sample: 1987 2019				
Included observations: 33				
Weighting series: W3				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	204.1818	2.622171	77.86748	0.0000
X2	0.966910	0.001507	641.6312	0.0000

X4	0.052592	0.000550	95.59155	0.0000
	Weighted Statistics			
R-squared	0.999969	Mean dependent var		2538.151
Adjusted R-squared	0.999967	S. d. dependent var		11571.35
S.E. of regression	2.969754	Akaike info criterion		5.101343
Sum squared reside	264.5832	Schwarz criterion		5.237390
Log likelihood	81.17217	Kerri Hannan - Quinn criter.		5.147119
F-statistic	482319.4	Durbin - Watson, stat		0.995528
Prob(F-statistic)	0.000000			
	Unweighted Statistics			
R-squared	0.999649	Mean dependent var		22220.32
Adjusted R-squared	0.999626	S. d. dependent var		31238.58
S.E. of regression	604.2153	The Sum squared reside		10952284
Durbin-Watson stat	1.226017			

WLS regression results were obtained according to the results

$$\hat{Y}_i = -204.1818 + 0.966910X2 + 0.052592X4$$

SE (2.622171) (0.001507) (0.000550)

T (77.86748) (641.6312) (95.59155)

R^2= 0.999969 = 482319.4, p = 0.000000 $\overline{R}^2 = 0.999967$ F

In order to analyze the correction of heteroscedasticity, after using WLS to estimate each model, the White test should be used to determine whether the model has heteroscedasticity again.

Table 16: White test results

Heteroskedasticity Test: White			
F-statistic	2.230328	Prob. F (4 p)	0.0912
Obs*R-squared	7.973804	Prob. Chi - Square (4)	0.0925
Scaled explained SS	10.54098	Prob. Chi - Square (4)	0.0322

Take the significance level =0.05, because of nR^2=7.973804<, so there is no heteroscedasticity. $\chi^2_{0.05} = 9.49$ W3 is the desirable weight.

Through comparison, it can be known that the weight W3 =1/ R2 has the best effect, and the revised model is:

$$\hat{Y}_i = -204.1818 + 0.966910X2 + 0.052592X4$$

SE (2.622171) (0.001507) (0.000550)

T (77.86748) (641.6312) (95.59155)

R^2= 0.999969 = 482319.4, p = 0.0000001 $\overline{R}^2 = 0.999967$ F

8. Autocorrelation test and correction

8.1 Autocorrelation test

The significance level of 1% is shown in the DW table, dL= 1.17 dU=1.29.The DW value of this model is 0.995528.Is an area that cannot be determined. At this point, we can only use the graphical method to check.

The residual figure of the model obtained by EViews software is shown below

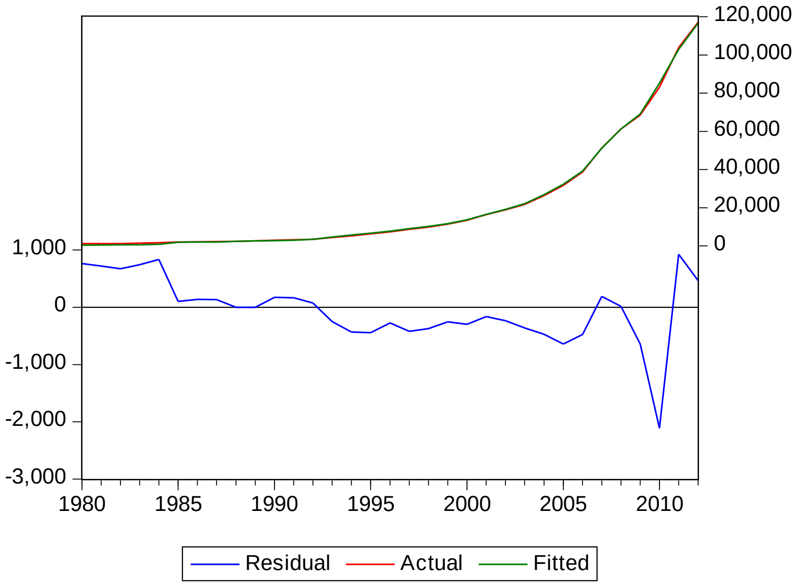


Figure 2: Residuals

It can be seen from the residual figure that the change of the residual has a systematic pattern, and the continuous is positive and the continuous is negative, indicating that there is a positive sequence correlation between the residual term. The conclusion of t statistic and F statistic in the model is not credible, so remedial measures should be taken.

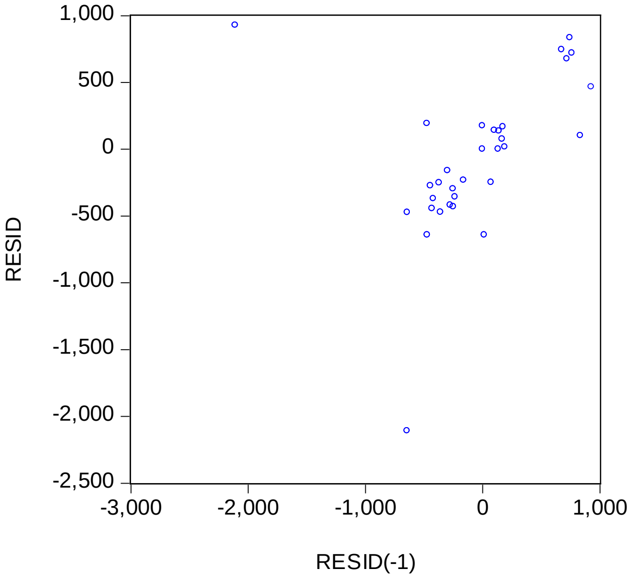


Figure 3: Residual scatter diagram

Residual test: From the scatter diagram of residual, it can be known that there may be a positive correlation between sequences.

The figure above shows that the model is clearly autocorrelated. Here we use the generalized difference method to compensate.

8.2 Autocorrelation correction

First, use the least square method to get the residual sequence E, select Quick\Generate Series from the main menu, enter E=resid in the dialog box, and click OK to get the residual sequence E. Use E for lag first order autoregression. Enter E E (-1) in the Eviews command bar to obtain the regression equation.

Table 17: OLS regression results

Dependent Variable: E		

Method: Least Squares				
Date: 06/20/20 Time: 20:18				
Sample (adjusted): 2007 2019				
Included observations: 32 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
E(-1)	0.357553	0.164429	2.174507	0.0374
R-squared	0.116318	Mean dependent var		76.66121
Adjusted R-squared	0.116318	S. d. dependent var		573.0903
S.E. of regression	538.7298	Akaike info criterion		15.44706
Sum squared resid	8997123.	Schwarz criterion		15.49286
Log likelihood	246.1529	Kerri Hannan - Quinn criter.		15.46224
Durbin-Watson stat	2.062834			

It can be seen that the generalized difference equation can be obtained by making generalized difference to the original model $\hat{\rho}=0.357553$

$Y_t-0.357553Y_{t-1}=\beta_0(1-0.357553)+\beta_1(X_2-0.357553X_{2-1})+\beta_2(X_4-0.357553X_{4-1})+\mu_t$ Enter y-0.357553 *Y (-1) C x2-0.357553 *X2 (-1) x4-0.357553 *X4 on the Eviews command bar, and enter the output

Table 18: OLS regression results

Dependent Variable: Y - 0.357553 * Y (1)				
Method: Least Squares				
Date: 06/20/20 Time: 20:25				
Sample (adjusted): 2007 2019				
Included observations: 32 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	292.3047	125.8524	2.322600	0.0274
X2-0.357553 * X2 (1)	1.057569	0.041361	25.56928	0.0000
X4-0.357553 * X4 (1)	0.027596	0.011217	2.460114	0.0201
R-squared	0.999495	Mean dependent var		15995.37
Adjusted R-squared	0.999460	S. d. dependent var		22024.66
S.E. of regression	511.8884	Akaike info criterion		15.40315
Sum squared resid	7598863.	Schwarz criterion		15.54056
Log likelihood	243.4504	Kerri Hannan - Quinn criter.		15.44870
F-statistic	28680.05	Durbin - Watson, stat		1.593213
Prob(F-statistic)	0.000000			

$$\hat{Y}_t^*=-292.3047+1.057569\hat{X}_{t2}^*+0.027596\hat{X}_{t4}^*$$

SE (125.8524) (0.041361) (0.041361)

T (-2.322600) (25.56928) (2.460114)

R^2 Df = 32 DW = = 0.999495 = 28680.05 $\overline{R}^2=0.999460$ 1.593213 F

Among them, $\hat{Y}_t^*=Y_t-0.357553Y_{t-1}$ $\hat{X}_2^*=X_2-0.357553X_{2-1}$ $\hat{X}_4^*=X_4-0.357553X_{4-1}$

Note: Due to the use of generalized differential data, the sample size is reduced by 1 to 32.The significance level of 5% can be seen from the DW statistical table, d_LD = 1.37_U=1.50, d in the $model_u$DW << 4 - d_u, so there is no autocorrelation.

Thus, the final financial revenue influencing factor model is obtained as follows:

$$\hat{Y}_t^*=-292.3047+1.057569\hat{X}_{t2}^*+0.027596\hat{X}_{t4}^*$$

SE (125.8524) (0.041361) (0.041361)

T (-2.322600) (25.56928) (2.460114)

R^2 Df = 32 DW = = 0.999495 = 28680.05 $\overline{R}^2 = 0.999460$ 1.593213 F

Among them, $\hat{Y}_t^* = Y_t - 0.357553 Y_{t-1}$ $\hat{X}_2^* = X_2 - 0.357553 X_{2-1}$ $\hat{X}_4^* = X_4 - 0.357553 X_{4-1}$

9. Conclusion and countermeasures

9.1 Conclusion

It can be seen from the model that in China, there is a high positive correlation between tax revenue and fiscal revenue, and the growth of tax revenue plays a significant role in promoting the growth of fiscal revenue.

The growth of fiscal revenue is strongly dependent on the fixed asset investment and has a positive correlation.

GDP (current price), the number of national employment on the fiscal revenue is not significant.

9.2 Countermeasures

(1)Strengthen tax collection and management, so as to ensure the steady growth of fiscal revenue. We will improve the tax structure, strengthen reform of the tax system, and make tax revenue more conducive to fiscal revenue growth.

(2)We should actively create a good investment environment, encourage domestic and foreign investment, and strengthen the role of investment as one of the troika in driving the economy and hence the fiscal revenue. At the same time, we should also change the economic development mode, change the existing investment-oriented economic development mode, and make economic development more dependent on consumption growth.

(3)We will improve the quality of the labor force and increase its contribution to government revenue. We will protect the rights and interests of workers, improve social redistribution, and reduce the large income gap.

(4)The growth of the national economy is the source of fiscal revenue growth, to develop the advanced productivity and enhance the level of residents' consumption and GDP boost national finance income increase, so that the country can have more money into the national infrastructure, boost GDP and the improvement of residents' consumption level, form a virtuous circle, practical and raise revenue.

(5)While raising fiscal revenue, we should pay attention to the combination of revenue and expenditure with fiscal expenditure, and make sure that fiscal revenue is used by the people. The government revenue should be used more effectively, more support should be given to industrial innovation, and more funds and resources should be invested in less developed areas to achieve social equity. It will promote the gross domestic product (GDP) of China's innovative industries, and at the same time promote the consumption level of Chinese residents, especially those in less developed areas with room for improvement, so as to realize the stable growth of fiscal revenue.

(6)Improve the quality and efficiency of economic operation and increase foreign trade. Normally, the ratio of the growth rate of fiscal revenue relative to GDP growth rate is called the elasticity coefficient of fiscal revenue to GDP, which is used to measure the dependent relationship between the growth rate of fiscal revenue and macroeconomic growth, showing how many percentage points of macroeconomic growth can drive the growth rate of fiscal revenue. There is no doubt that this index is of great significance for the government to implement macro-control, adjust industrial structure and improve the quality of economic operation in real economic work. The increase of import and export will lead to the prosperity of China's international trade, thus stimulating fiscal growth. Therefore, it is necessary to promote the development of import and export trade and increase the trade cooperation between countries and regions.

(7) Determine scientific fiscal and tax policies to control the proportion of fiscal revenue to GDP within a reasonable range. The proportion of fiscal revenue to GDP is a very important index to measure the government's control over the national economy. The higher the proportion of fiscal revenue to GDP, the more capable the country is to provide citizens with abundant public services.

However, this target must be controlled within a reasonable range. If the proportion is too high and the financial resources concentrated by the government are excessive, it will crowd out the interests of taxpayers, weaken the foundation of economic development, and ultimately affect the development of national economy and the growth of fiscal revenue. If the proportion is too low, the government's concentrated financial resources will be limited, which will seriously affect the normal performance of various government functions, and weaken the fiscal control ability of macroeconomic operation and optimal allocation of resources.

References

[1] CAI Wei, Peng Gang.China's industrial industry from the perspective of spatial econometrics R&Research on Spillover Effect [J]. Journal of Donghua University of Technology (Social Science edition), 2008,37(04):323-327.
[2] Li Jinhua. Frontiers and New Challenges of Quantitative Economics [N]. China Social Science Journal,2017-08-30(004).
[3] National Bureau of Statistics.China Statistical Yearbook 2010 [M]. Beijing: China Statistical Press, 2019.
[4] Ma Haitao. China's Tax system [M]. Beijing: China Renmin University Press, 2019.
[5] Zhang Xiaodong. Econometric Analysis [M]. Beijing: Economic Science Press, 2013.
[6] Chen Gong. Finance [M].Beijing: Renmin University of China Press, 2017
National Bureau of Statistics of the People's Republic of China. China Statistical Yearbook 2011 [M].Beijing: China Statistics Press, 2011.
[8] Liu Wei, Chen Zhao. (in Chinese)Econometrics software Eviews 8.0 Modeling Methods and operational skills [M].Beijing: Machinery Industry Press, 2019
[9] Liu Wei, Chen Zhao.Econometrics software Eviews 5.0 Modeling methods and operation Techniques [M].Beijing: Machinery Industry Press, 2019
[10] Tan Yun. Tax Compliance, Tax Service and Optimization of Tax collection and administration efficiency in China [J].Journal of Zhongnan University of Economics and Law,2018, (6).
[11] Li Yudong, Gu Zhenhua. (in Chinese)Research on applied Course Reform of Business Economics -- A Case study of Business Data Analysis Course [J].Journal of Hubei University of Economics (Humanities and Social Sciences edition),20,17(06):152-154.
Shang Chao. Research on coordinated Economic Development of Counties in Central Plains Economic Zone from the perspective of high-quality Development -- Based on Spatial Econometrics [J].Business Economics,2020(06):20-25.
[13] Kang Shuju.Spatial econometric analysis and exploration of China's economic growth [J]. Value engineering,2019,38(29):216-217.
[14]Qiu Junping, Shen Paid Chen, SONG Yanhui.Progress and Trends in Econometrics research at home and abroad in the past decade -- A Visual comparative Study based on Citespace [J].Modern intelligence,2019,39(02):26-37.
[15] Plum Nai, Pan Wenqing.Econometrics [M]. Beijing: Higher Education Press, 2005.
[16] Gao Hongye.Western Economics [M]. Beijing: Renmin University of China Press, 2017.