Study on the Impact of Industrial Structure on GDP and Economic Growth in China Based on Multiple Regression

Liang Xiaoyi
Beijing Jiaotong University, School of Economics and Management, Beijing, China
liangxiaoyi_17@163.com

Key Words：Industrial Structure, GDP, Economic Growth, Regression Model

Abstract. The transformation of industrial structure and economic growth are interdependent and mutually reinforcing. Under certain conditions, the change of industrial structure is the basis of economic growth and the main factor to promote economic growth.This paper summarizes the GDP and the growth rate of three major industries in China from 1992 to 2016.By establishing a multiple linear regression model, this paper studies the contribution of the growth of three major industries to China's economic growth, and concludes the importance of adjusting the industrial structure in transforming the mode of economic development and promoting the sustainable development of China's economy.

1. Research Summary

Chenery(1960) analyzed the empirical data of 51 countries to show that when a country's economic scale changes, the service industry and agriculture change the least, while the manufacturing industry grows the most. Therefore, he proposed a model of industrial growth, and believed that this model of industrialization can optimize the allocation of resources.Kuznets(1949) compared empirical data from 50 countries and found that the growth of the manufacturing sector will be accompanied by the growth of per capita national income.

In recent years, Chinese scholars have done some research on China's industrial structure and economic growth. Most scholars use cointegration theory and Granger causality test theory to conduct empirical analysis to study the relationship between economic growth and industrial structure. The conclusion is that before 2002, China's economic growth was mainly driven by institutional reform and tertiary industry.In recent years, scholars believe that secondary and tertiary industries play an important role in promoting economic growth.

2. Data Collection

Table 1 shows the GDP of China from 1992 to 2016.
The growth rate can be used as an index to analyze the impact of various industries on GDP. By calculating and sorting out the data in Table 1, we can get Table 2, which is the summary of annual growth rate from 1993-2016.

Table 1 Table of cumulative GDP 1992-2016

Data Source：http://www.stats.gov.cn/

Unit: 100 million yuan				
Year	GDP	Primary industry	Secondary industry	Tertiary industry
1992	26923	5867	11700	9357
1993	35334	6964	16454	11916
1994	48198	9573	22445	16180
1995	60794	12136	28679	19978
1996	71177	14015	33835	23326
1997	78973	14442	37543	26988
1998	84402	14818	39004	30580
1999	89677	14770	41034	33873
2000	99215	14945	45556	38714
2001	109655	15781	49512	44362
2002	120333	16537	53897	49899
2003	135823	17382	62436	56005
2004	159878	21413	73904	64561
2005	184937	22420	87598	74919
2006	216314	24040	103720	88555
2007	265810	28627	125831	111352
2008	314045	33702	149003	131340
2009	340903	35226	157639	148038
2010	397983	40497	186481	171005
2011	489301	46163	227039	216099
2012	540367	50902	244643	244822
2013	595244	55329	261956	277959
2014	643974	58344	277572	308059
2015	689052	60862	282040	346150
2016	744127	63671	296236	384221

Table 2 The Growth Rate of GDP and Industries 1993-2016

Unit: %				
Year	Growth Rate of GDP	Growth Rate of Primary Industry	Growth Rate of Secondary Industry	Growth rate of Tertiary industry
1993	31.24	18.70	40.64	27.34
1994	36.41	37.46	36.41	35.78
1995	26.13	26.78	27.77	23.48
1996	17.08	15.49	17.98	16.76
1997	10.95	3.04	10.96	15.70
1998	6.87	2.60	3.89	13.31
1999	6.25	-0.32	5.20	10.77
2000	10.64	1.18	11.02	14.29
2001	10.52	5.60	8.68	14.59
2002	9.74	4.79	8.86	12.48
2003	12.87	5.11	15.84	12.24
2004	17.71	23.19	18.37	15.28
2005	15.67	4.70	18.53	16.04
2006	16.97	7.23	18.40	18.20
2007	22.88	19.08	21.32	25.74
2008	18.15	17.73	18.42	17.95
2009	8.55	4.52	5.80	12.71
2010	16.74	14.96	18.30	15.51
2011	22.95	13.99	21.75	26.37
2012	10.44	10.27	7.75	13.29
2013	10.16	8.70	7.08	13.54
2014	8.19	5.45	5.96	10.83
2015	7.00	4.32	1.61	12.36
2016	7.99	4.61	5.03	11.00

3. Model Establishment

The model was established according to the statistical data collected from 1993 to 2016. Its expression is as follows:

$$Y = \alpha_0 + \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \varepsilon \qquad (1)$$

Y represents the annual growth rate of gross domestic product (GDP). X_1, X_2 and X_3 represent the annual growth rate of the first, second and third industries respectively. α_0 represents the inherent growth rate of the economy under constant conditions.It can be approximated that GDP growth is the weighted growth rate of three industrial added value, while α_i represents the weight of each industrial sector in economic growth, and $\alpha_i X_i$ represents the contribution of each industrial sector

to economic growth. ε represents a random error term.

Through formula (1), we can understand how the GDP will change with each industry increasing by 1 percentage point, so as to make economic forecast and provide basis and reference for industrial policy adjustment.

4. Model Checking

The regression analysis results can be obtained through multiple linear regression.

Dependent Variable: Y

Method: Least Squares

Date: 12/15/18 Time: 22:16

Sample: 1993 2016

Included observations: 24

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.042989	0.298920	0.139521	0.8921
X1	0.171006	0.013988	12.05986	0.0000
X2	0.435275	0.020814	23.95960	0.0000
X3	0.387501	0.028536	13.59264	0.0000

R-squared	0.998586	Mean dependent var	15.90897
Adjusted R-squared	0.998263	S.D. dependent var	8.307018

S.E. of regression	0.337915	Akaike info criterion	0.859605
Sum squared resid	1.598130	Schwarz criterion	1.058531
Log likelihood	-3.746158	F-statistic	3418.992
Durbin-Watson stat	1.160573	Prob(F-statistic)	0.000000

4.1 Goodness of fit test

The results of regression parameter estimation show that the sample determinant R^2 = 0.9986 and the adjusted determinant coefficient is 0.9983, which shows that the model fits the sample well.

4.2 F-test

For H_0: $\alpha_1=\alpha_2=\alpha_3=0$, given the significance level α=0.05, the critical values $F_\alpha(3,14)=3.34$ with degrees of freedom k-1=3 and n-k=14 were found in the F distribution table. From the OLS regression analysis table, F = 3418.992. Because F = 3418.992 >$F_\alpha(3,14)$ = 3.34, the original hypothesis should be rejected, indicating that the regression equation is significant, that is, "primary industry", "secondary industry", "tertiary industry" and other variables combined do have a significant impact on GDP.

4.3 T-test

For H_0: $\alpha_j=0$（j=1、2、3、4）, given a significant level $\alpha = 0.05$, $t_{\alpha/2}(n-k)$ = 2.145 when the degree of freedom is n-k=14. The t-statistics corresponding to α_0， α_1， α_2 and α_3 are 0.139521, 12.05986, 23.95960 and 13.59264 respectively. Therefore, the t-test of α_0 is less than $t_{\alpha/2}(n-k)$, and its t-test is not significant, but the coefficient of determination of the model is high, and the F-test value is also significant, which indicates that there may be multiple collinearity.

4.4 Multiple collinearity test

Table 3 Correlation matrix

	X_1	X_2	X_3
X_1	1.000000	0.810689	0.841092
X_2	0.810689	1.000000	0.882356
X_3	0.841092	0.882356	1.000000

Table 4 Univariate Regression Estimation Results

Variable	X_1	X_2	X_3
Parameter Estimates	0.7157	0.8016	1.2228
T-Statistics	8.1950	16.7191	12.0997
R^2	0.8076	0.9459	0.9015
Adjusted R^2	0.7956	0.9425	0.8953

The correlation coefficients among the explanatory variables are relatively high, which proves that there may indeed be multiple collinearities.

Stepwise regression is used to test and solve multiple collinearity problems. The unitary regression of Y to X_1, X_2 and X_3 is made, and the results are shown in Table 4.

The adjusted R^2 of the equation with X_2 is the largest. Therefore, on the basis of X_2, stepwise regression with other variables was added. The results are shown in Table 5 .

Table 5 Regression results with adding new variables

	X_1	X_2	X_3	Adjusted R^2
X_1，X_2	0.2575 （5.47130）	0.5859 （12.0288）		0.9795
X_2，X_3		0.4995 （9.5015）	0.5378 （6.5463）	0.9841

When X_1 and X_3 were added, the adjusted R^2 increased and t-test was significant. The possibility of col-linearity may be reduced by choosing variables as relative numbers. The final regression results are as follows:

$$Y = 0.0436 + 0.1704X_1 + 0.4411X_2 + 0.3876X_3 \qquad (2)$$

4.5 Heteroscedasticity test

Using White test, the results show that nR^2 = 0.605477*18 = 10.898586 < the critical value 16.9190, so there is no heteroscedasticity in the model.

4.6 Sequence correlation test

DW=1.1606, given the significance level $\alpha = 0.05$, check the Durbin-Watson table, n = 18, k = 3, and get the lower critical value d_L= 0.933, d_U= 1.696, d_L< DW = 1.1606 < d_U. According to the judgment region, it can not be determined whether there is autocorrelation or not.
The following results are obtained by using Cochrane-Orcutt iteration method.

Dependent Variable: Y
Method: Least Squares
Date: 12/16/18 Time: 18:35
Sample(adjusted): 1994 2016
Included observations: 23 after adjusting endpoints
Convergence achieved after 6 iterations

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.357593	0.286210	-1.283019	0.2241
X1	0.143857	0.012425	11.61430	0.0000
X2	0.478103	0.019608	24.60288	0.0000
X3	0.397866	0.025741	16.01067	0.0000
AR(1)	0.252992	0.197086	1.288093	0.2222

R-squared	0.999306	Mean dependent var	15.53771
Adjusted R-squared	0.998892	S.D. dependent var	7.665621
S.E. of regression	0.247069	Akaike info criterion	0.290598
Sum squared resid	0.738617	Schwarz criterion	0.535661
Log likelihood	2.529923	F-statistic	3810.285
Durbin-Watson stat	1.707125	Prob(F-statistic)	0.000000

Inverted AR Roots	.25

After correction, DW = 1.7071, when n = 17, k = 3, get d_L= 0.897, d_U= 1.710. At this time, the value of DW is larger than that of d_L, but close to the value of d_U. Although the autocorrelation has not been completely eliminated, we think that the value of DW in this model is better than that of above-mentioned DW = 1.1606. The revised regression equation is as follows:

$$Y = -0.357593 + 0.1439X_1 + 0.4781X_2 + 0.3979X_3 \qquad (3)$$

5. Conclusions and suggestions

According to the formula (3), when the primary industry, the secondary industry and the tertiary industry increase by 1 percentage point, China's economy increases by 0.1443, 0.4782 and 0.3986 percentage points, respectively.

Therefore, from the regression data and related tests, we can get the relationship between the three major industries and China's GDP growth. The conclusion is that at present, the secondary industry contributes the most to economic growth, followed by the tertiary industry and the primary industry.

The following policy suggestions are drawn. Industry, especially heavy industry, has always been the pillar of China's economic development. There are differences in the internal structure and regional structure of industry. Emphasis is laid on upgrading the manufacturing industry, fostering and developing strategic emerging industries, establishing an innovative country, and improving China's core competitiveness.

Efforts should be made to create and maintain a market environment conducive to fair competition and promote enterprises to optimize and upgrade their industrial structure in accordance with market mechanism; actively promote the establishment of a unified national market system; through deepening reform, abandon local protectionism and market blockade, avoid convergence of industrial structure, and build a unified national market system.

References

[1] S. Kuznets, National Income and Industrial Structure, *Econometrics*, vol.17, pp. 205-241, 1949.

[2] H. B. Chenery, Patterns of Industrial Growth, *The American Economic Review*, vol.50, pp. 624-654, 1960.

[3] J. Ke, The relationship between industrial structure upgrading and economic growth, *Statistics and Decision-making,* vol.11, 2008.

[4] H. Sun, Research on the Influencing Factors and Changes of Industrial Structure Height in China, *Population & Economics*, vol.3, 2012.