Journal of Southeast University (English Edition) Vol. 25, No. 4, pp. 541 – 544 Dec. 2009 ISSN 1003—7985

Forecasting of development of the Jiangsu construction industry and its case analysis

Lu Yan[1] Li Junna[2]

([1]School of Civil Engineering, Southeast University, Nanjing 210096, China)
([2]Zhengzhou Track Transportation Company Ltd., Zhengzhou 450008, China)

Abstract: In order to grasp the development path of the Jiangsu construction industry, a multivariable linear regression model for forecasting is proposed. Five factors affecting development of the Jiangsu construction industry are chosen as explanatory variables. They are the construction industry's fixed assets K, the gross domestic product (GDP), real estate added value (REAV), construction industry export (WS) and investment in construction and installation projects (JA). The principal component analysis is used to resolve multicollinearity between them. The construction added value (CAV) is chosen as a dependant variable, and the growth model of the Jiangsu construction industry is established. Statistical data from 1990 to 2008 are used to test the prediction accuracy of the model. The predictive results show that from 2009 to 2012, the average annual growth rate of the Jiangsu construction industry added value will be 17.65% while the GDP growth rate will be 14.16%; the Jiangsu construction industry will grow faster than the GDP in the near future. The construction output of the GDP continues to rise, and its pillar position will be further strengthened.

Key words: construction industry; development; forecasting; principal component analysis

As a pillar industry of the national economy, the Jiangsu construction industry not only makes great contribution for national economy, but also has high degree of correlation with other industries. Therefore, studying the growth of the Jiangsu construction industry and predicting the trend of its development can help to guide the formulation of relevant policies and can ensure its healthy development.

In the aforementioned research, the main factors affecting the Jiangsu construction industry have been determined including the construction industry's fixed assets K, the gross domestic product (GDP), real estate added value (REAV), construction industry export (WS) and investment in construction and installation projects (JA). The construction added value (CAV) is chosen as a dependant variable and the five variables above are employed as explanatory variables. By using the principal component analysis to resolve multicollinearity, the growth model is built to forecast the development of the Jiangsu construction industry.

1 Research Method

1.1 Multicollinearity test

Multicollinearity[1] is a frequent problem in multiple linear regression models. According to the causality test results from the aforementioned research, a strong correlation exists in the factors affecting the development of the Jiangsu construction industry. Therefore, when using these five explanatory variables to build the growth model of the Jiangsu construction industry, the multicollinearity test should be conducted first to ensure that the model not only better fits but is also a true reflection of the economic growth of the Jiangsu construction industry in actual situations. Using the software Eviews[2–3], the simple correlation coefficient of each variable is calculated to test the multicollinearity of the model.

1.2 Principal component analysis

The new combined variables obtained from the original variables by the principal component analysis[4–6] are alternatives to the original variables. As the results of the combination with the original variables, they play an important part in overcoming the overlapped information caused by multicollinearity, and they also can eliminate the effects to regression modeling caused by multicollinearity.

2 Analysis of Factors Affecting Development of Jiangsu Construction Industry

2.1 Results of multicollinearity test

The correlation coefficient test of ln$K$, lnGDP, lnREAV, lnWS, and lnJA (using statistical data from 1990 to 2008[7]) is carried out by "corr" order of Eviews 5.0. We can obtain the results as shown in Tab. 1.

Tab. 1 Correlation coefficients of factors affecting development of Jiangsu construction industry

Factors	lnCAV	ln$K$	lnGDP	lnREAV	lnWS	lnJA
lnCAV	1.000	0.967	0.991	0.997	0.986	0.961
ln$K$	0.967	1.000	0.978	0.972	0.949	0.980
lnGDP	0.991	0.978	1.000	0.995	0.991	0.985
lnREAV	0.997	0.972	0.995	1.000	0.989	0.972
lnWS	0.986	0.949	0.991	0.989	1.000	0.960
lnJA	0.961	0.980	0.985	0.972	0.960	1.000

The correlation coefficient test shows that the correlation coefficient between the variables is relatively great; therefore, the model built by OLS has serious multicollinearity. In this paper, the principal component regression is used to resolve multicollinearity.

2.2 Results of principal component analysis

Principal component analysis of five factors affecting the development of the Jiangsu construction industry is carried out by SPSS16.0[8–9]. The standard Z-score formula is

Received 2009-07-21.
Biography: Lu Yan(1978—), female, doctor, lecturer, luyanseu@126.com.
Foundation item: The Program of the Construction Department of Jiangsu Province(No. JS2007-17).
Citation: Lu Yan, Li Junna. Forecasting of development of Jiangsu construction industry and its case analysis[J]. Journal of Southeast University (English Edition), 2009, 25(4): 541 – 544.

used.

First, the KMO test and Bartlett's test of sphericity are carried out. The results show that these five variables are suitable for factor analysis and are not independent of each other, so the principal component analysis can be carried out. Tab. 2 shows the results.

Tab. 3 shows that the eigenvalues of the first principal component are greater than 1, and the cumulative is 98.153%, so this principal component can be extracted.

Tab. 2 Results of KMO test and Bartlett's test of sphericity

KMO	Approximate chi-squre	Degree-of-freedom	Significance
0.805	223.628	10	0.000

The coefficients in the factor load matrix of each factor are in a descending order. It is clear that the load distribution of each factor in the principal component is greater than 0.985 as shown in Tab. 4.

Tab. 3 Results of variance contribution analysis

Component	Initial eigenvalues			Extraction sums of squared loadings		
	Total	Variance/%	Cumulative/%	Total	Variance/%	Cumulative/%
1	4.908	98.153	98.153	4.908	98.153	98.153
2	0.063	1.260	99.413			
3	0.021	0.419	99.833			
4	0.007	0.139	99.972			
5	0.001	0.028	100.000			

Tab. 4 Factor load matrix

Influencing factors	Component load
lnGDP	0.999
lnREAV	0.995
lnJA	0.988
lnWS	0.987
ln$K$	0.985

According to the regression coefficient matrix of the principal component score as shown in Tab. 5, we can draw the relationship among the principal components as

Tab. 5 Factor score matrix of principal component

Influencing factors	Component load
lnGDP	0.204
lnREAV	0.203
lnJA	0.201
lnWS	0.201
ln$K$	0.201

$$F_1 = 0.201Z\ln K + 0.204Z\ln\mathrm{GDP} + 0.203Z\ln\mathrm{REAV} + 0.201Z\ln\mathrm{WS} + 0.201Z\ln\mathrm{JA} \quad (1)$$

3 Forecasting of Development of Jiangsu Construction Industry

3.1 Relationship between principal component and development of construction industry

Regression of principal component F_1 and ZlnCAV is advanced by the least-squares method. The stationary test and co-integration analysis are carried out to dependent variable ZlnCAV and explanatory variable F_1. The test standard is the SIC information criterion. We can obtain the following results of the unit root test for both the components F_1 and ZlnCAV by Eviews 5.0 as shown in Tab. 6.

From the results of the unit root test, we can see that the ZlnCAV and F_1 are both second-order unit root series. Their co-integration equation is carried out by OLS, and then the unit root test is carried out on to residual series E_8 of the co-integration equation. The results are shown in Tab. 7.

Tab. 6 Unit root test results of dependent variable ZlnCAV and explanatory variable F_1

Sequence	Regression form	Statistic t	Critical value		
			1%	5%	10%
F_1	(1, 1, 3)	−3.355 594	−4.886 426	−3.828 975	−3.361 984
$F_1(-1)$	(1, 0, 0)	−2.048 686	−3.959 148	−3.081 002	−2.681 330
$F_1(-2)$	(0, 0, 0)	−3.681 562	−2.740 613	−1.968 430	−1.604 392
ZlnCAV	(1, 1, 3)	−2.623 389	−4.886 426	−3.828 975	−3.362 984
ZlnCAV(−1)	(1, 0, 0)	−2.409 578	−3.959 148	−3.081 002	−2.681 330
ZlnCAV(−2)	(0, 0, 0)	−7.773 039	−2.740 613	−1.968 430	−1.604 392

Tab. 7 Unit root test results of residual series E_8

Sequence	Regression form	Statistic t	Critical value		
			1%	5%	10%
E_8	(0, 0, 1)	−2.260 219	−2.728 252	−1.966 27	−1.605 026

From Tab. 7, we can see that the residual series in the 5% significance level reject the assumption of having a unit root; therefore, it is a smooth sequence and it is a co-integration between ZlnCAV and F_1. But the constant term of the co-integration equation directly built by OLS is not significant. After canceling the constant term and using regression, we can obtain

$$Z\ln\mathrm{CAV} = 0.989\,511F_1 \quad (2)$$
$$27.399\,96$$
$$(0.000\,0)$$
$$R^2 = 0.979\,133,\ D.W = 0.383\,278$$

The coefficient in Eq. (2) gets through the 1% significant level test, except for *D. W* is equal to 0.383 278, and it is far less than 2, which shows that there is autocorrelation[10-11] in the model. By using the generalized difference method, we obtain

$$\begin{gathered} Z\ln CAV = 0.839\ 911 F_1 \\ 27.399\ 96 \\ (0.000\ 0) \\ R^2 = 0.510\ 908,\ D.\ W = 1.551\ 187 \end{gathered} \tag{3}$$

The coefficient in Eq. (3) gets through the significant level test, and *D. W* is equal to 1.551187, which is larger than 1.34, so it gets through the 10% significant level test, and it shows that there is no autocorrelation[10-11] in the model. Eq. (3) reflects the long-term equilibrium relationship between principal component F_1 and lnCAV.

3.2 Building the growth model of Jiangsu construction industry

Substituting Eq. (1) into Eq. (3), we can obtain

$$\begin{aligned} Z\ln CAV &= 0.839\ 911(0.201 Z\ln K + 0.204 Z\ln GDP + \\ &\quad 0.203 Z\ln REAV + 0.201 Z\ln WS + 0.201 Z\ln JA) = \\ &\quad 0.168\ 822 Z\ln K + 0.171\ 342 Z\ln GDP + \\ &\quad 0.170\ 502 \ln REAV + 0.168\ 822 Z\ln WS + \\ &\quad 0.168\ 822 Z\ln JA \end{aligned} \tag{4}$$

Standardized data can be converted to the original data by the following equation:

$$\left.\begin{aligned} b_i &= \frac{S_y}{S_i}\beta_i \qquad i = 1, 2, \ldots, k \\ b_0 &= \bar{y} - \sum_{i=1}^{k} b_i \bar{x}_i \end{aligned}\right\} \tag{5}$$

where y and x_i are dependent variables and explanatory variables, respectively; S_y and S_i are the standard deviations of y and x_i, respectively; β_i is the coefficient of explanatory variables; b_0 and b_i are the constant term and the coefficient of explanatory variables of the converted model, respectively. The mean value and the standard deviation of these variables can be obtained by the Descriptives order of SPSS16.0. Then we can obtain the results as shown in Tab. 8.

Tab. 8 Mean value and standard deviation of lnCAV and main influencing factors

Variables	Sample quantity	Minimum	Maximum	Constant term	Standard deviation
lnCAV	19	4.07	6.58	5.524 7	0.804 79
ln*K*	19	2.93	5.27	3.971 7	0.781 09
lnGDP	19	7.26	9.32	8.326 4	0.627 20
lnREAV	19	3.30	5.92	4.739 2	0.812 66
lnWS	19	3.98	7.31	5.817 2	0.992 45
lnJA	19	5.65	7.76	6.603 5	0.647 95

b_i and b_0 can be calculated from the results in Tab. 8, and we can obtain the growth model of the Jiangsu construction industry by

$$\begin{aligned} \ln C\hat{A}V &= 0.320\ 1 + 0.173\ 9\ \ln K + 0.171\ 3 \ln GDP + \\ &\quad 0.136\ 9 \ln REAV + 0.175\ 6 \ln WS + \\ &\quad 0.214\ 5 \ln JA \end{aligned} \tag{6}$$

The coefficients of the model are in line with the economic significance. The fitting prediction diagram of the forecasting model can be obtained by the PLOT order of Eviews 5.0, as shown in Fig. 1.

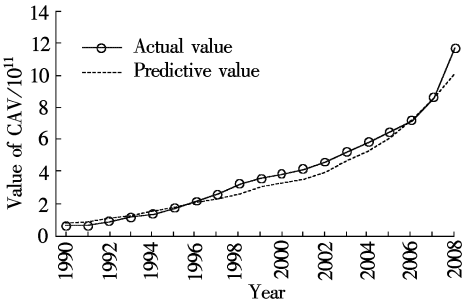

Fig. 1 The fitting prediction of the growth model of Jiangsu construction industry

By intuitive judgment, we can see that the predictive value and the actual value are very fitting, but we should test the prediction accuracy of the model. Percentage error (PE), mean percentage error (MPE) and mean absolute percentage error (MAPE) are the indices commonly used to test the prediction accuracy. According to the model, we can obtain the predictive values and the actual values of CAV and PE of the Jiangsu construction industry from 1990 to 2008, as shown in Tab. 9.

Tab. 9 Predictive values and actual values of CAV and PE of Jiangsu construction industry

Year	Actual value	Predictive value	PE/%
1990	58.460 0	84.549 2	-44.63
1991	64.550 4	93.819 6	-45.34
1992	92.601 4	107.296 2	-15.87
1993	115.701 2	129.903 2	-12.27
1994	135.001 4	154.685 2	-14.58
1995	168.821 3	180.397 2	-6.86
1996	213.252 6	209.507 4	1.76
1997	264.073 0	235.542 0	10.80
1998	322.812 8	262.952 6	18.54
1999	356.813 7	302.897 5	15.11
2000	383.114 3	328.315 3	14.30
2001	411.079 3	354.146 1	13.85
2002	460.820 1	396.372 0	13.99
2003	521.648 1	467.564 1	10.37
2004	585.286 0	528.216 3	9.75
2005	647.681 0	609.881 7	5.84
2006	716.982 7	730.525 1	-1.89
2007	864.063 5	859.490 9	0.53
2008	1 166.893 0	1 011.224 0	9.89

According to the data in Tab. 9, we can obtain MPE = -0.88%, MAPE = 14.01%; that is to say, the prediction ability of this model is strong.

3.3 Predictive results of development of Jiangsu construction industry

According to the Eleventh Five-Year Plan of Jiangsu Province, the GDP of the whole province will be about 2 900 billion Yuan, with an annual average growth rate of more than 10%. By the statistical data, the GDP growth rate of the Jiangsu province in 2005, 2006, 2007 and 2008 are 14.75%, 14.48%, 14.9% and 12.5%, and we take

the mean value as 14.16%. In accordance with the estimate of the Eleventh Five-Year Plan, the added value of the real estate industry will be 130 billion Yuan in 2010, while the number was 91.478 billion in 2006, so the average annual growth rate of the added value of the real estate industry is 9.18% in the Eleventh Five-Year Plan. The Eleventh Five-Year Plan has no relationship to the growth plan of the net fixed asset and the external output of the construction industry as well as the construction and installation investment, so we select the average cycle-increase rate of these indices in the recent five years for forecasting. We can obtain the following results as shown in Tab. 10.

Tab. 10 Growth rate of each factor used for forecasting %

Index	K	GDP	REAV	WS	JA
Growth rate	19.19	14.16	9.18	37.93	20.51

According to the growth rate in Tab. 10, the prediction of the next four years of the Jiangsu construction industry is obtained by the growth model, and the results are shown in Tab. 11.

Tab. 11 Predictive values of CAV of Jiangsu construction industry

Year	Predictive value of CAV	Predictive value of GDP	Construction output of GDP/%
2009	1 189.744	16 770.485	7.09
2010	1 399.780	19 237.423	7.28
2011	1 646.894	22 067.248	7.46
2012	1 937.635	25 313.340	7.65

The growth rate of the CAV of the Jiangsu construction industry is 17.65% calculated by the predictive values in Tab. 11, which shows that the Jiangsu construction industry also has a tremendous development potential. And the construction output of the GDP continues to rise, which shows that its pillar position will be further strengthened in future years.

4 Conclusion

In this paper, the growth model of the Jiangsu construction industry is established, and the statistical data from 1990 to 2008 are used to test the prediction accuracy of the model. It shows that the prediction ability of this model is strong. The prediction of the Jiangsu construction industry from 2009 to 2012 is carried out by the growth model.

The predictive results show that the Jiangsu construction industry will grow faster than the GDP and the construction output of the GDP continues to rise during the next four years. So relevant policies and management measures of the Jiangsu construction industry should support its continuous stability and development. Then the Jiangsu construction industry can develop rapidly in the future.

References

[1] Mishra S K. Multicollinearity and maximum entropy leuven estimator[J]. *Economics Bulletin*, 2004, **25**(3): 1-11.

[2] Griffiths William E, Hill R Carter, Lim Guay C. *Using Eviews for principles of econometrics*[M]. New York: Wiley, 2008: 10-150.

[3] Gao Huixuan. *Application of multivariate statistics*[M]. Beijing: Beijing University Press, 2005: 25-40. (in Chinese)

[4] Dunteman George H. *Principal components analysis*[M]. London: Sage Publications, 2005: 10-45.

[5] Schökoph B, Smola A, Müller K. Kernel principal component analysis [C]//*Advances in Kernel Methods*: *Support Vector Learning*. Cambridge: MIT Press, 1999: 327-352.

[6] Hotelling H. Analysis of a complex of statistical variables into principal components[J]. *Journal of Educational Psychology*, 1993, **24**(3): 417-444.

[7] The Statistics Bureau of Jiangsu Province. *Jiangsu statistical year-book* [M]. Beijing: China Statistics Press, 1990-2008. (in Chinese)

[8] Field Andy P. *Discovering statistics using SPSS*[M]. London: Sage Publications, 2005.

[9] Santner T J, Duffy D E. *The statistical analysis of discrete data* [M]. New York: Springer-Verlag, 1989: 142-194.

[10] Ding C, Helleseth T, Lam K Y. Several classes of binary sequences with three-level autocorrelation [J]. *IEEE Trans Information Theory*, 1999, **45**(7): 2606-2612.

[11] Ding C, Helleseth T, Martinsen H M. New families of binary sequences with optimal three-level autocorrelation[J]. *IEEE Trans Information Theory*, 2001, **47**(1): 428-433.

江苏省建筑业增长预测及实证分析

陆 彦[1] 李俊娜[2]

([1] 东南大学土木工程学院, 南京 210096)

([2] 郑州市轨道交通有限公司, 郑州 450008)

摘要: 为了把握江苏省建筑业的发展轨迹, 提出了其增长预测的多元线性回归模型. 选择影响江苏省建筑业增长的5个因素, 即建筑业固定资产 K、地区生产总值GDP、房地产增加值REAV、建筑业对外输出WS和建筑安装工程投资JA作为解释变量, 使用主成分分析解决其多重共线性问题. 再选定建筑业增加值lnCAV作为被解释变量, 建立江苏省建筑业增长模型, 并用1990~2008年的统计数据检验模型预测精度. 预测结果表明: 2009~2012年江苏建筑业增加值的平均增长率将为17.65%, 同期GDP的增长率为14.16%, 江苏省建筑业未来增速快于GDP增速. 建筑业产值占GDP比重持续上升, 其支柱地位将逐步加强.

关键词: 建筑业; 增长; 预测; 主成分分析

中图分类号: F407.9