Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/51/20230621

Research on the Relationship Between Higher Education Development and Economic Growth in China

Bingcheng Guo[1,a,*]

[1]*School of business, Macau University of Science and Technology, Macau, 999078, China*
a. b4110118@163.com
**corresponding author*

Abstract: Education has always been a very important research issue, and China's higher education has developed rapidly along with its economic take-off. The issue of the correlation between education and economic development has also been researched many times in the academic world. So, this paper analyzes the dynamic correlation between the development of higher education and China's economic development, as well as the gross domestic product of the People's Republic of China (excluding Hong Kong Special Administrative Region of China, Macao Special Administrative Region of China, and Taiwan Province of China) for the period 1978-2021, based on the existing literature, through the data on the number of graduates from higher education, general higher education, and postgraduates (master's degree and doctorate). It is finally concluded that economic growth will directly contribute to the development of higher education, and similarly, the growth of higher education will offset economic growth with a slight lag.

Keywords: higher education, economic growth, human capital, VAR model, China

1. Introduction

Education has always been one of the hot issues of public concern, and higher education is an important part of education. Since the Reform and Opening-up in 1978, higher education in China has experienced rapid development and popularization, and its impact on economic development has become one of the focuses of attention in both academic and policy circles. It is generally accepted that higher education is beneficial to economic development, and most policymakers are clearly aware of this, with most regions seeing higher education as an important way to promote economic development. As an important source of technological innovation, higher education institutions (universities, colleges, etc.) not only contribute technological capital to society but also provide a large number of high-quality human resources. Previous studies have also pointed out that education, as a major component of human capital, has a significant impact on the economy [1-2]. This paper will use econometric methods to develop a VAR model of the relationship between China's tertiary education development level and economic growth and, at the same time, use the granger causality test, impulse response function, and analysis of variance (ANOVA) to investigate the dynamic correlation between the two. The analyses in this paper can provide significant references for future economic and educational policies, not only in China, but also in various other countries or regions.

DOI: 10.54254/2754-1169/51/20230621

2. Review of Empirical Literature

Few previous studies on the impact of education on economic growth have been analyzed quantitatively using econometric models. “Higher education has a statistically significant positive impact only on developed countries, while primary education has a positive impact on less developed countries, and secondary education has a positive impact on developing countries”, describing the different impacts of different levels of education on different types of countries [3]. Citing the Cobb-Douglas Function to study the correlation between higher education institutions and GDP per capita, it was obtained that "significant differences in GDP per capita between European (EU) countries are associated with the activities of higher education institutions [4]. There are also articles based on the VAR model to study higher education and positive growth, but the data chosen is really regional such as Fujian Province or Chongqing Municipality, and there is little analysis of the national data for the time being [5-6].

3. Empirical Analysis

3.1. Modelling

3.1.1. Introduction to the VAR Model

Vector Autoregressive Models (VAR) are modeled based on the statistical nature of the data, and the VAR model is constructed by treating each endogenous variable in the system as a function of the lagged values of all the endogenous variables in the system. The VAR model is commonly used to forecast interconnected time-series systems and to analyse the dynamic shocks of stochastic perturbations to systems of variables, thus explaining the impact of various economic shocks on the formation of economic variables.

The expression for a p-order VAR model without exogenous variables and with k endogenous variables is:

$$y_t = \beta_1 y_{t-1} + \beta_2 y_{t-2} + \cdots + \beta_p y_{t-p} + \varepsilon_t \qquad t = 1,2,\dots\dots,T \tag{1}$$

Where: $yt=(y_{1t}, y_{2t}\dots y_{kt})'t=1, 2,\dots\dots,T$; p is the lag order; β_j, j=1, 2,…, p is the k*k order parameter matrix; ε_t is the k-dimensional white noise vector, which can be contemporaneously correlated with each other, but not with their own lagged values and not with the variables on the right side of the equation [7].

3.1.2. Data and Variance

In this paper, annual data from 1978 to 2021 were selected. The data comes from the Bloomberg database, CSMAR database and the National Bureau of Statistics of China [8-10]. In this paper, the sum of the number of general college graduates and the number of postgraduate graduates (highedu) from 1978 to 2021 is considered an indicator of higher education development, while the gross domestic product (GDP) of China from 1978 to 2021 is considered an indicator of economic development. However, this paper takes the natural logarithm of the two variables to eliminate the heteroskedasticity in the data, and the variables after taking the natural logarithm are expressed as lnhighedu and lngdp.

DOI: 10.54254/2754-1169/51/20230621

3.1.3. Model Estimation

3.1.3.1. Unit Root Test

The smoothness of the variables needs to be tested before performing the regression analysis. If the variables selected are not smooth, then the regression that results from the analysis of the time series is not a regression in the true sense of the word, but a pseudo-regression.

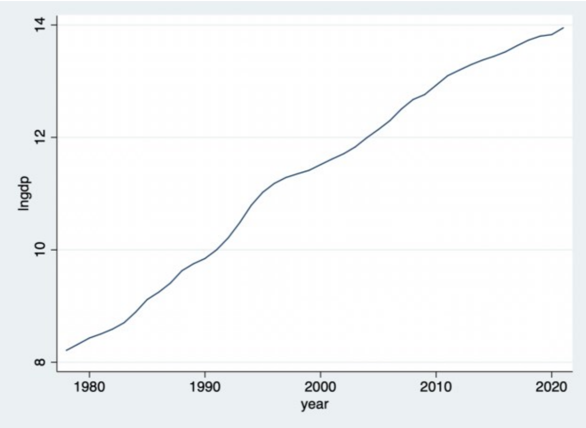

Figure 1: Trend graph of lngdp 1978-2021.

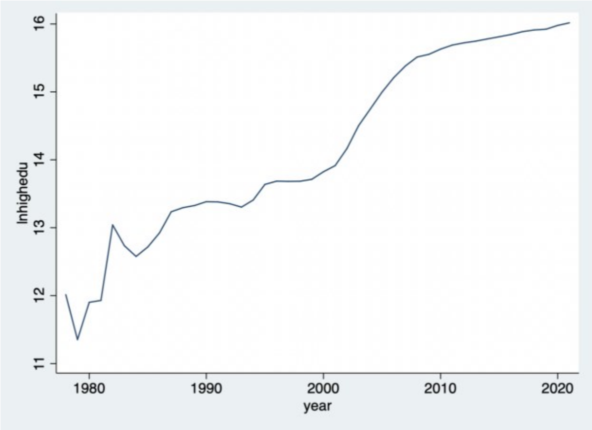

Figure 2: Trend graph of lnhighedu 1978-2021.

Figures 1 and 2 give the information that these two variables are not smooth. So ADF (Augmented Dickey-Fuller) was chosen to test the unit root of the regression in this paper. As shown in Table1, the author takes a significance level of 10%.

Table 1: Unit root test.

Variables	Test statistic	1% critical value	5% critical value	10% critical value	results
lngdp	-2.012	-3.628	-2.950	-2.608	not stationary
lnhighedu	-0.900	-3.628	-2.950	-2.608	not stationary
d.lngdp	-2.817*	-3.634	-2.952	-2.610	stationary
d.lnhighedu	-9.152***	-3.634	-2.952	-2.610	stationary

Notes: standard errors in parentheses *p<0.1 **p<0.05 ***p<0.01, d.lngdp denotes the first order difference of lngdp and d.lnhighedu denotes the first order difference of lnhighedu. The author obtained these results by Stata14.

From the above table, both lngdp and lnhighedu are I (1), the first order differences of each of the variables lngdp and lnhighedu are smooth.

3.1.3.2. Establish Model

Since the first-order differences of each of lngdp and lnhighedu are smooth, a VAR model can be established. In this paper, the lag order chosen is 2 according to the required criteria of the FPE, HQIC, and SBIC guidelines, so the model can be built as:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t \quad (2)$$

where y_t=(lngdp$_t$, lnhighedu$_t$)’ t=1978,……, 2021;

The VAR(2) model coefficient regression is shown in Table 2.

Table 2: Model coefficient estimation results.

lngdp	Coef.	Robust Std.	t-stat	p-value	[95% Conf. Interval]	
lnhighedu	.007935	.0248795	0.32	0.752	-.058393	.0425229
lngdp L1.	1.577564	.1377262	11.45	0.000	1.298243	1.856886
lnhighedu	.0095718	.0200293	0.48	0.636	-.0310496	.0501932
lngdp L2.	-.6105308	.1369836	-4.46	0.000	-.8883464	-.3327151
lnhighedu	.0354353	.0221648	1.60	0.119	-.009517	.0803876
β_0	-.0965547	.1358554	-0.71	0.482	-.3720821	.1789728

Notes: L1. L2. denote the first-order lag and the second-order lag, respectively. we get these results by Stata14.

Table 2 shows that the coefficients of the vector autoregressive models lngdp L1. lngdp L2. are extremely significant, i.e., the first-order after of lngdp has an extremely significant positive effect on lngdp, and the second-order lag of lngdp has a significantly smaller negative effect on lngdp. The coefficients of lngdp L2. are very close to the 10 percent level of significance.

3.2. Model Testing

3.2.1. Results of the Test

Table 3 below shows the results of the test for the lag order of the VAR (2) model.

Table 3: Test results for the lag orders of the model.

	Equation: lngdp			Equation: lnhighedu			Equation: All		
lag	chi2	df	p-value	chi2	df	p-value	chi2	df	p-value
1	189.651	2	0.000	28.2479	2	0.000	218.272	4	0.000
2	32.0839	2	0.000	1.11243	2	0.573	33.1518	4	0.000

Notes: chi2 denotes Chi-squared test statistic, df denotes degrees of freedom. These are Wald test results for lag order of VAR(2) model gotten by Stata14.

lngdp is particularly significant in all lag orders. lnhighedu is significant in first order lag while second order lag is not significant. But two variables, lngdp and lnhighedu are jointly significant in each lag order.

DOI: 10.54254/2754-1169/51/20230621

3.2.2. Goodness of Fit Test

The goodness of fit of the two equations among the VAR(2) model is .9994 and .9785, respectively (as shown in Table 4). This responds to the fact that the explanatory power of the equations is extremely strong.

Table 4: Goodness of fit test.

Equation	RMSE	R-sq	chi2	P>chi2
lngdp	.044585	.9994	71574.04	0.0000
lnhighedu	.196762	.9785	1908.268	0.0000

3.2.3. Robustness Test

Figure 3 shows the AR root diagram in the VAR (2) model. All points are located in the unit circle, so it can be judged that the VAR model is stable, so pulse response analyses and variance contribution analyses can be performed on the basis of this model.

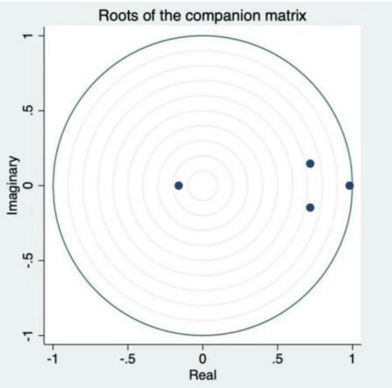

Figure 3: AR root diagram in the VAR (2) model.

3.3. Granger Causality Test

The Granger causality test can help explore the mutual causality between economic development and higher education. The author examines the granger causality test statistic for each lag and finds that, at the 10% level of significance, lngdp granger causes lnhighedu at lag one, while lnhighedu does not cause lngdp. After lag two lngdp granger cause lnhighedu and lnhighedu granger cause lngdp as well (The granger causality test results are shown in Table 5.) It indicates that although the development of higher education has a certain lag in the promotion of economic growth, the lag is very limited, and economic growth is bound to have a very positive effect on the development of higher education.

Table 5: Granger causality test results.

Lag length	Null hypothesis	Obs	chi2	Prob>chi2
1	lnhighedu does not granger cause lngdp	43	1.7435	0.187
	lngdp does not granger cause lnhighedu		4.8628	0.027
2	lnhighedu does not granger cause lngdp	42	4.0049	0.045
	lngdp does not granger cause lnhighedu		7.8357	0.005
3	lnhighedu does not granger cause lngdp	41	5.4104	0.020
	lngdp does not granger cause lnhighedu		13.212	0.000

DOI: 10.54254/2754-1169/51/20230621

Table 5: (continued).

4	lnhighedu does not granger cause lngdp	40	5.6051	0.018
	lngdp does not granger cause lnhighedu		12.583	0.000

3.4. Impulse Response Analysis and Variance Decomposition

In time series analysis, impulse response functions (IRFs) play a crucial role in understanding the dynamics between variables in a system. The impulse response function is usually associated with the study of linear time-invariant systems and describes the behaviour of the system over time after the introduction of an impulse, revealing the response to change. The impulse response analysis of the model provides a more intuitive understanding of the short-term and long-term interactions between education and economic development (Impulse response functions for higher education and economic development are plotted using Stata.)

From Figure 4, we can see that after giving the impact of one shock in the current period of higher education, higher education itself reacts faster to itself; it drops rapidly from the initial 0.22 and gradually stabilizes after the fifth period, with the impulse response stabilizing at around 0.05. This suggests that education itself has continuity and that its ongoing development is a virtuous circle. But a current shock of education for economic development (GDP) shows a slow rising trend, from the initial 0 value to 0.03 in the fifth period after the leveling off. This shows that China's higher education has a positive effect on the development of the economy, but its impact has a lag; the results are in line with the objective law, that is, the higher education itself, for the long time needed to cultivate talents as well as talents back to the community, has a certain time lag.

When a current shock impact of economic development (GDP) is given, economic development itself has a long-lasting and stable impact, initially decreasing from 0.06 at the very beginning to 0.05 in the twelfth period, which indicates that the country's economy continues to grow and is in an upward phase. A current shock to economic development (GDP) will make it possible to raise the level of education, specifically from the earliest period to the fourth period of rapid growth, from a value of 0 to 0.018, showing a rapid upward trend, and converging to 0.02 in the fifth period and thereafter, reflecting that economic development for the promotion of the level of education is very obvious and that the development of the economy can be reflected in the raising of the level of education more quickly.

To sum up, it takes about five years for the improvement of education levels to play a stable and obvious role in promoting economic development. However, economic development can respond more quickly to the improvement of education levels, and the two interact in both directions.

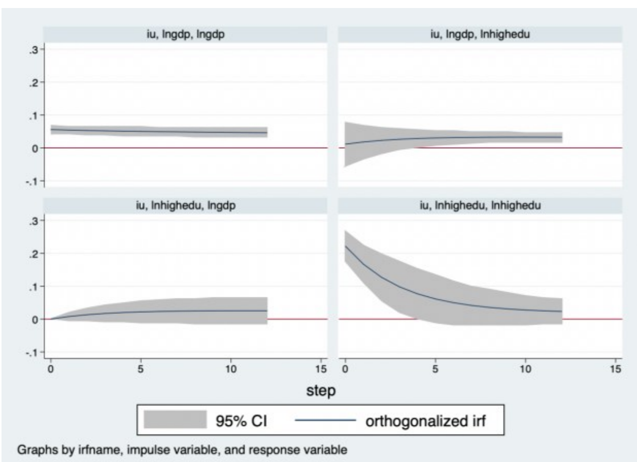

Figure 4: Impulse response analysis.

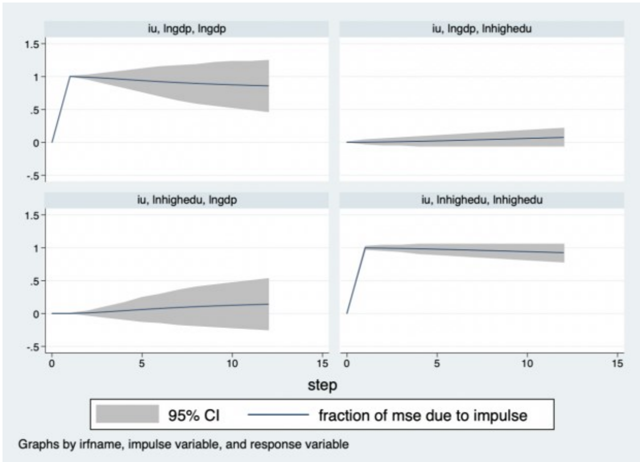

Figure 5: Variance decomposition.

Analysis of variance (ANOVA) is a statistical technique that essentially quantifies the total variance explained by different factors, which helps to measure the actual significance of the observed effects.

The ANOVA chart drawn by Stata (as shown in Figure 5) shows that both the education level and economic development in China are affected by their own influence and gradually decrease over time, but to a great extent are still dominated by their own influence. From the ANOVA chart of economic development, we can see that the fluctuating influence of economic development itself gradually decreases from 100% at the very beginning to 70%, which indicates that our economic development model has good stability, and in the long run, economic development is also influenced by many other factors. At the same time, the contribution rate of the fluctuation of education level to the fluctuation of economic development is not significant in the initial periods, from 0 in the initial period to 3.3% in the fifth period, and then grows rapidly to 17.6% in the twelfth period after the fifth period, which demonstrates that the enhancement of the level of education has a limited contribution to the economic development in the initial period, but with the passage of time, the effect is gradually revealed, and in the long run education has a great contribution of education to economic development in the long run.

The ANOVA chart of education level shows that the fluctuation of education level is mainly affected by itself, and its variance contribution decreases over time, but the trend of decrease is not obvious, from 100% at the beginning to 92% in the 12th period. Economic development also plays a role in the volatility of education, with the contribution of fluctuations in economic development to fluctuations in education rising gradually from 0 in the initial period to 9 percent, suggesting that economic development is also an important factor affecting the level of education.

4. Conclusion

In this paper, the data from 1978 to 2021 are chosen to build a lagged two-order vector autoregressive model, VAR (2), on the number of higher education graduates and GDP. Through the granger causality test, impulse response analysis, and variance decomposition, it can be learned that economic growth will directly promote the development of higher education, and likewise, the development of higher education will counteract economic growth with a slight lag. However, this paper analyzes only a single variable, rather than multiple variables, yielding only correlational conclusions. The study did not prove a deeper, direct causal relationship between high education and economic growth. The author will also build a more complex and comprehensive model and manipulate multiple variables in subsequent studies, observing the effects of these manipulations on other variables to arrive at more convincing and deeper causal relationships.

References

[1] *Lucas Jr, R. E. (1988). On the mechanics of economic development. Journal of monetary economics, 22(1), 3-42.*
[2] *Romer, P. M. (1990). Capital, labor, and productivity. Brookings papers on economic activity. Microeconomics, 1990, 337-367.*
[3] *Mingat, A., & Tan, J. P. (1996). The full social returns to education: Estimates based on countries' economic growth performance.*
[4] *Pastor, J. M., Peraita, C., Serrano, L., & Soler, Á. (2018). Higher education institutions, economic growth and GDP per capita in European Union countries. European Planning Studies, 26(8), 1616-1637.*
[5] *Lu, Chengxiao. (2012). Study on the relationship between higher education scale and economic growth in Fujian--an empirical study based on VAR model. Journal of Longyan College, 30(6), 46-50.*
[6] *Sun, C. (2020). Effect of higher education on GDP in Chongqing based on VAR model. International Journal of Social Science and Education Research, 3(4), 52-59.*
[7] *Stock, J. H., & Watson, M. W. (2015). Introduction to econometrics (3rd updated edition). Age (X3), 3(0.22).*
[8] *Data Source. 1978-2021. Bloomberg. http://www.bloomberg.com*
[9] *Data Source. 1978-2021. CSMAR. https://cn.gtadata.com*
[10] *Data Source. 1978-2021. National Bureau of Statistics. http://www.stats.gov.cn*