An Empirical Study on the Relationship between Educational Expenditure and the Average Wage of Employed People in China

Cui Yang

Ulanqab Vocational College, Ulanqab City, Inner Mongolia Autonomous Region, China

Abstract: With the reform and opening up and socialist modernization, China's economy has achieved rapid growth. In 2012, the financial input to education accounted for 4% of the total financial input of the state. Education is related to people's livelihood, according to the theory of human capital, investment in education is conducive to the promotion of personal income. In this paper, the VAR model is constructed, impulse response analysis and variance decomposition analysis are carried out, and the relationship between educational investment and individual income is discussed.

1. Introduction

With the reform and opening up and the establishment of a socialist market economy, China has achieved high economic growth. However, since the 1990s, income inequality in China has gradually increased, which has become one of the causes of social conflicts in China. There are many factors leading to income inequality, among which the unequal distribution of educational resources is an important one. The distribution of educational resources has many meanings, including the input of educational funds. As early as 1960, after the "human capital theory" was first proposed by the American economist Theodore, the human capital theory was founded by Schulz, and then developed till now, people have extensively studied the impact of educational funds input on individual income.

2. Literature Review

At present, there are many researches on the relationship between educational expenditure and income gap at home and abroad. It is widely acknowledged in the academic circle that there is an inverted U-shaped relationship between education quality and income gap [1]. According to the discovery of scholar Ana, there is an average number of years to receive education, which is generally seven years. The Kuznets curve has an inflection point, and the threshold effect is similar to the inverted "U" shaped relationship [2]. Tang Qingjie estimated that the threshold value of educational expansion on income inequality was 7.7294 [3]. Zhang Wen et al believe that the relationship between educational investment and income gap is non-linear[4]. Some scholars hold that there are two influencing relationships between education and income: Omona believes that in different social political and economic fields, there should be both "structure effect" and "wage compression effect" in education, but its influence on income imbalance is still uncertain at present [5]. Li Zhuo, Lei Xin

and Gong Feng argue that through education and other efforts, the income of people born in different generations changes differently. People born around 1960 and 1970 have the most drastic income decline, those born in the 1980s have the second largest income decline, and those born in the 1950s have the least impact [6].

3. Research Review

3.1 Human Capital Theory

The term "human capital" was first proposed by American economist Theodore Roosevelt in 1960. The theory of human capital originated from economic studies. It was founded by economist Schulz and further developed by Becker. This theory opens up a new way of thinking about human's understanding of economics. Human knowledge and skills, as well as human's own value and physical quality are all counted as human beauty capital. The theory of human capital mainly focuses on the improvement of production efficiency brought by the improvement of human knowledge and skills[7]. Human capital is mainly characterized by externality, accumulation and expansion [8]. The outward expansion of the possible boundary of production in society is determined by these characteristics. Capital contains human resources, which will attract individuals and society to invest in this kind of capital, and it is clear that the return rate for this kind of investment will be higher and higher [9].

According to the human capital theory, the increase of investment in education will promote the improvement of education level, so as to improve individual productivity, obtain greater economic value and obtain higher personal income. When the level of personal income increases, more attention will be paid to the education of the next generation, that is, the "intergenerational transmission" of educational ideas, and the education level of the next generation will also be improved, thus obtaining a positive cycle[10]. Figure 1 could show the theory:
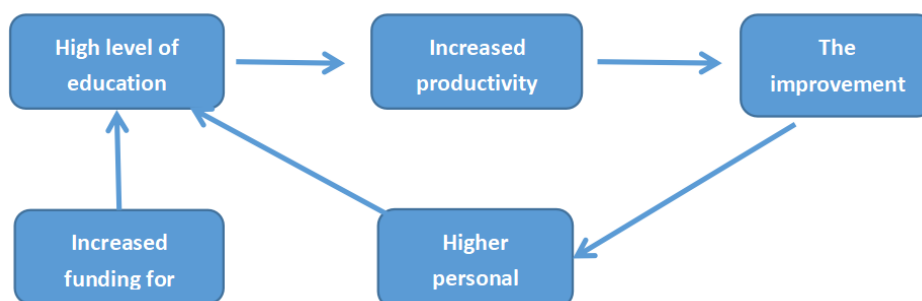

Figure 1: Greater emphasis on educating the next generation

3.2 Vector Autoregressive Model

In econometrics, vector autoregressive model is often used. It is often abbreviated as VAR. Vector autoregression has been around for a long time, dating back to the 1980s. People construct vector autoregression models mainly for the following considerations: The time series analysis was expanded from a single time series data to multivariate time series. At any t time interval, the observed samples changed from 1 to N, The time series also expanded from single to multivariate; Standard Autoregressive models (AR) with simple expressions could not work well for multivariate time series analysis[11].

In order to study how the economic impact brought by the investment of national education funds had an impact on personal income, this paper selected the investment of education funds from 2009 to 2017 as the independent variable, and the average wage of urban private unit employees as the dependent variable to construct the VAR model for analysis and testing. In terms of the selection of

dependent variables, it is believed that the average wage of urban private unit employees can better reflect the market orientation.

4. Model and Design

4.1 Vector Autoregressive Model Test

In fact, compared with a single time series data, multivariate time series data is more common and is composed of a single time series. For multivariate time series data [12], vector autoregression model adopts a more flexible time series modeling strategy: Given multivariate time series data, for any T-th time interval, the following linear expression exists:

$$Y_t = \varepsilon + A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + \varepsilon_t \ (t = 1,2,\cdots,T)$$

Where, $Y_t = (y_{1,t}, y_{2,t}, \cdots y_{N,t})^t$ $\varepsilon = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_N)^t$; for A_i : And $i = 1,2,\ldots,P \varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, \ldots\ldots, \varepsilon_{N,t})^t$; Where, Y_t is a vector of endogenous variables of order Nx 1. ε is a column vector of constant terms of Nx1 order, $Y_{t-i}(i = 1,2,\ldots\ldots,p)$ is the endogenous variable vector of lag; A_p is the NxN dimensional matrix of the estimated parameters, P is the endogenous lag order, ε_t is a vector composed of N dimensional random error term, its elements can be related with each other, but cannot be associated with the lag term and the variables on the right side of its expression.

4.2 Data Source, Variable Setting and Processing

Through consulting China Statistical Yearbook, we selected and sorted out the input of China's education funds and the average salary of urban private employees from 2009 to 2017 as indicators. After natural logarithm processing and heteroscedasticity elimination, the two variables were denoted as LNY and LNX, respectively. This could be shown in Table 1.

Table 1: Statistical table of educational fund input and average salary of employed personnel

Year	Educational expenditure	Average wages
2009	165027065	18199
2010	195618471	20759
2011	238692936	24556
2012	286553052	28752
2013	303647182	32706
2014	328064609	36390
2015	361291927	39589
2016	388883850	42833
2017	425620069	45761

Source: China Statistical Yearbook
Note: The average salary is the average salary of employed persons in urban private units

4.3 Establish Var Model

4.3.1 Test the Stationarity of Data

ADF test is usually used to further test the robustness of the data. According to Table 2, the test results of the two variables in the table indicate that, at the significance level of 5%, the ADF test value is less than the critical value, which means that the null hypothesis is rejected, they have no unit

root, and the sequence is a stable sequence of first-order single integers.

Table 2: Results of ADF test

Variables	ADF test value	5% threshold at the significant level	P value
LNX	3.369993	3.320969	0.0469
LNY	6.746919	3.403313	0.0016

Data source: Eviews analysis results

4.3.2 Determine the Order of Lag

As can be seen from the above, the data passed the stationarity test, and the vector autoregressive model can be established on this basis. However, before modeling, it is necessary to first determine the hysteresis order of the model. Too short a lag period will prevent the dynamic changes of the model from being reflected more accurately. Too long a lag period will affect the predictive effect of the model. In this paper, statistics and statistical criteria as shown in Table 3 are used to determine the order of lag.

Table 3: Results of hysteresis test

Lag period	LogL	LR	FPE	AIC	SC	HQ
0	17.58628	NA	4.01 e-05	4.453223	4.468678	4.644235
1	48.31318	3511646 *	217e-08*	12.08948 *	12.13584	12.66252
2	52.23525	2.241184	4.07 e-08	12.06722	12.14449 *	13.02227 *

Data source: Eviews analysis results

According to the results determined in Table 3, among the five criteria, FPE criterion, AIC criterion and HQ criterion all set the optimal lag time as 1, so the optimal lag order of the VAR model finally determined is 1.

4.3.3 Test of Co-Integration Relation

Under the classical linear regression model, such as the multiple regression model, in order to make the statistical results unbiased and consistent (that is, with the wireless increase of the sample, the estimated value is infinitely close to the real value), multiple requirements are put forward for the model, among which the dependent variable and the independent variable are required to be stationary time series. Non-equilibrium stability makes the co-integration based on non-stationary series and the accompanying error correction model more and more widely used. For several non-stationary time series, if the linear combination variables composed of them have the characteristics of stationary, then the variables can be determined to have a co-integration relationship, which can be interpreted as the long-term equilibrium relationship between these variables[13]. To further explain the relationship between educational expenditure and average individual wages, that is, the long-term co-integration relationship between an equilibrium relationship, it is necessary to reintegrate the data to test its co-integration relationship. According to the ADF test above, these two time series meet the conditions of co-integration test: non-stationary first-order single-order series meet the test conditions. Johanson's co-integration test is applied below, and the results show that the co-integration equation exists at the 5% level of the time series. The co-integration equation is cointegrated as follows:

$$LNY\ LNX - 1.018761 = 9.505979 \tag{1}$$

$$(0.040097)$$

According to the equation, the coefficient of elasticity of educational input to individual average

wage growth is 1.018761 and the standard deviation is 0.040097. According to the standard deviation, it can be judged that LNY and LNX have a co-integration relationship and the co-integration relationship is relatively stable. From a long-term perspective, every 1% increase in educational investment can lead to 1.018761% increase in individual average wages.

4.3.4. Establishment of Var Model

Since the original sequence is stable at the same stage, co-integration test can be performed, and the co-integration test proves that the sequence is co-integration, and a VECM model (error correction model) can be established on the original sequence. The VECM model obtained is as follows:

$$
\begin{bmatrix} LNY \\ LNX \end{bmatrix}
= \begin{bmatrix} 19.27180 \\ 1.327106 \end{bmatrix} + \begin{bmatrix} -0.097641 & 0.064834 \\ 2.403105 & 0.991668 \end{bmatrix} \begin{bmatrix} LNY \\ LNX \end{bmatrix}_{t-1} + \begin{bmatrix} 2.403105 & 0.991668 \\ -0.574619 & -0.126119 \end{bmatrix} \begin{bmatrix} LNY \\ LNX \end{bmatrix}_{t-2} + \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \end{bmatrix}
$$

$$\text{Where t=2009, 2010...... 2017} \tag{2}$$

5. Data Analysis

In order to observe the degree of fit of VAR model and the relationship between education funds input and the average wage of urban private employees, robustness test, impulse response analysis and variance decomposition analysis should be carried out on the model.

5.1 Robustness Test

Robustness test aims to test whether the relationship between the core independent variable and the dependent variable is robust. If the coefficient in front of the core variable is stable, the result can be considered robust. The robustness of the model is tested, and the results shown in Figure 2 are as follows
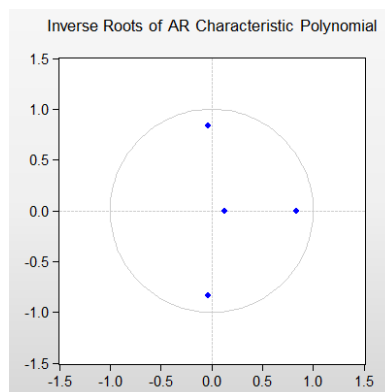

Data source: Eviews analysis results

Figure 2: Root and unit circle graphs of AR characteristic polynomials in VAR model

As shown in Figure 1, there are altogether 4 feature roots of the model established, all of which fall within the unit circle above. The above results show that the established VAR model is stable, and the next step of impulse response analysis can be carried out for the stable model.

5.2 Impulse Response Analysis

Impulse response analysis is an important step in econometric analysis using vector autoregressive model. Their primary purpose is to describe the evolution of the impact of a model variable on one or more variables. This makes them a very useful tool in evaluating the economy, and is the impact of one standard deviation shock of a random disturbance term on system dynamics [14]. Impulse impact analysis describes the impact of the change of one endogenous variable in the VAR model on the current and future values of other endogenous variables. Figure 3 below shows the graph obtained by impulse response function analysis:
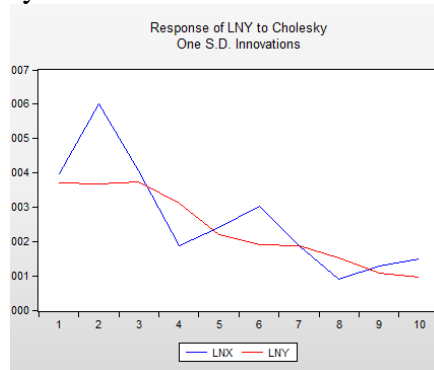
Note: The x-coordinate in the figure represents the number of tracking lag periods of the impulse response function, and the y-coordinate represents the impulse response function of the dot line table of response degree

Figure 3: Impulse response of LNY to LNX impact

As can be seen from the picture, LNX's impact on LNY showed a shock downward trend during the period 1-8, and an upward trend during the period 8-10, while LNY's impact on LNY itself showed a wave downward state. In general, the investment in education funds had a huge impact on individual average wages in all periods.

5.3 Analysis Of Variance Decomposition

The principle of variance analysis is to decompose the total discretization of test data into discretization of different factors, and make data estimation, so as to find the importance of each factor in the total discretization. Through variance decomposition analysis, it is possible to estimate how much influence the edge drawing situation of independent variables and dependent variables can have on the VAR model and the effect of the number of lag periods obtained above[15].
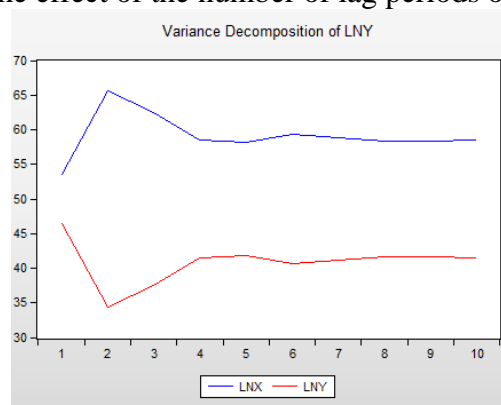

Figure 4: Analysis of variance

As can be seen from Figure 4, LNX's interpretation rate of LNY is over 50%, much higher than that of LNY itself. In the 1-2 periods, the interpretation rate reached the highest, and then dropped sharply to about 57%. In the 8-10 periods, it remained at about 57%, indicating that the investment in education has an influence on personal income and the influence is stable.

6. Conclusion

From 2004 to 2017, the number of regular college students in China increased from 4.47 million to 7.61 million. The number of graduate students has also continued to increase, which has led to an overall improvement in the quality of the population[16]. As can be seen from the changing trend of enrollment, China's higher education is moving towards a more comprehensive talent cultivation model.

The increase in the number of students enrolled in higher education has also led to a significant increase in the enrollment rate of higher education (estimated based on the number of students enrolled and the number of people born 18 years ago). In 2017, the undergraduate enrollment rate exceeded 20 percent, an increase of 17 percentage points since 2000. The number of regular institutions of higher learning reached 2,631 in 2017, an increase of nearly 1,500 over 2000. Although both the higher education institutions and the number of students have increased significantly, the ratio of higher education students to students has increased since 2000, due to the relatively slow growth in the number of teachers. Teacher training has become a key breakthrough point for the continued rapid development of higher education. The development of China's education is not only reflected in the basic popularization of 9-year compulsory education, but also in the rapid growth of high school and higher education enrollment rate. On average, the proportion of students in every 100,000 people is basically maintained at about 20,000, among which the proportion of students in higher education is constantly increasing. At the same time, the proportion of illiterate people over the age of 15 in the population has decreased to less than 5%[16].

From the trend of China's education development in the past 10 years, it is not difficult to see that higher education is the fastest developing field in China's education, whether the number of enrollment, enrollment rate, the number of teachers, the number of schools have maintained a rapid development trend. Secondary (high school, middle school) and primary education, affected by the birth population, enrollment number, the number of schools have declined. The enrollment rate and the number of teachers in senior high schools have maintained growth [17]. However, the popularity of nine-year compulsory education has exceeded 90%, and the number of teachers grows relatively slowly. With the improvement of the enrollment rate of higher education, the proportion of the population with higher education in the population above 6 years old will gradually increase[18].

So does the investment in education really promote the development of human capital? By sorting out the relationship between education investment and individual wage income from 2009 to 2017, a VAR model was established, and the mutual influence, long-term and dynamic relationship between the two was observed and analyzed. The following conclusions can be drawn:

First, based on the co-integration relationship test of VAR model, it is found that there is an equilibrium relationship between educational funds input and personnel wage growth, which can be regarded as stable in the long run. From the long-term relationship, every 1% increase in educational expenditure can lead to 1.018761% increase in personal income. Educational investment has the characteristics of non-profit, indirect, continuous, delayed, cyclical and so on. These characteristics determine that government investment is the main source of educational funds, which requires the government to make long-term investment.

Secondly, from the perspective of impulse response analysis and variance analysis, education investment accounts for most of the impact on the increase of personal income. In addition, the

increase of individual average salary income can also explain part of the increase in salary, which can be explained by the selection and adjustment of the market itself under prosperous social and economic conditions. The advantages of human capital will gradually emerge with the passage of time, so the investment in education funds should be long-term and continuous. The results of VAR model are used for empirical analysis. The results of this paper show that there is a positive interaction between the national financial investment in education funds and the average salary of urban private units. We should be careful to deal with the problems in the development process, speed up the pace of development, and realize the win-win situation of education and economic development.

References

[1] Liu Yunbo, Yin Lu & Guo Jianru. (2020). The quality of higher education and overeducation: Where should higher education funding go? The Finance Research Letters (prepublish). Doi: 10. 1016, j. f rl. 2020. 101824.

[2] Ana Balcao Reis. (2020). Public funding of higher education: Who gains, who loses? International Journal of Economic Theory(2). doi:10. 1111, ijet. 12176.

[3] Tang Qingjie. (2019). An empirical analysis of the influence of educational years on personal income. Think Tank Time (43), 213-214+216.

[4] Zhang Wen, Mei Jun & Wang Lulu. (2021). Analysis on regional differences and influencing factors of public investment performance in higher education in China. Heilongjiang Higher Education Research (01), 66-72. doi: 10. 19903, j. cnki. cn23-1074/g. 2021. 01. 012.

[5] J Omona. (2012). Funding Higher Education in Uganda Modalities Challenges and Opportunities in the twenty first Century. Makerere Journal of Higher Education.

[6] Gong Feng & Lei Xin. (2010). Quantitative measures of Chinese style fiscal decentralization. Statistical Research (10), 47-55. doi:10. 19343, j. cnki. 11-1302/c. 2010. 10. 008.

[7] Francesco Andreoli & Alessio Fusco. (2019). Application of the EU-SILC 2011 data module "intergenerational transmission of data. Disadvantage "to robust analysis of inequality of opportunity. The Data in Brief, (C), doi: 10. 1016, j. ib. 2019. 104301.

[8] Shen Z Y. (2020). Research on rural anti-poverty problem in western China from the perspective of human capital theory. Rural Science and Technology (29), 34-35. doi:10. 19345, j. cnki. 1674-7909. 2020. 29. 021.

[9] Hao W W, Zhou H T. (2000). The development of secondary vocational education and the reform of higher normal education in less developed areas: A survey and analysis of vocational middle schools in four provinces. Educational Research (04), 35-41.

[10] Sedat ALATA Xian & Mesut CAKIR. (2016). The Effect of Human Capital on Economic Growth: A Panel Data Analysis. Yonetim Bilimleri Dergisi (27).

[11] Bhatti M. Ishaq & Al Shanfari Hatem (2017). Econometric Analysis of Model Selection and Model Testing. Taylor and Francis.

[12] Cai Wenbo & Cheng Lu. (2015). An empirical analysis of the relationship between economic growth and the scale of higher education in Xinjiang based on VAR model. Heilongjiang Higher Education Research (12), 52-56.

[13] Mina Toksoz. (2017). Wages of oil, parliaments and economic development in Kuwait and the UAE/Business politics in the Middle East/Transformation of the Gulf: politics, economics and the global order, After the Sheikhs, The "coming collapse of the Gulf monarchies. Middle Eastern Studies (2), doi: 10. 1080/00263206. 2016. 1213245.

[14] Li Huajun & Liu Si. (2021). Higher Education, Technological Innovation, and Regional Sustainable Development: Insights from a VAR Model, Discrete 'Dynamics in Nature and Society, doi: 10. 1155/2021/8434528.

[15] Zhao S K, Yu H Q & Liu Z L. (2011). Theoretical model and empirical research on the relationship between investment in higher education and economic growth. China Higher Education Research (09), 11-15. doi: 10. 16298, j. cnki. 1004-3667. 2011. 09. 002.

[16] Yan Meng & An Zhen. (2021). The impacts of high frequency US uncertainty shocks on China's investment and bank loans: Yan Meng & An Zhen. (2021). Evidence from mixed - frequency VAR. Applied Economics Letters (1), doi: 10. 1080/ 13504851. 2020. 1725232.

[17] Su Hui & Sun Yi. (2016). An empirical study on the correlation between secondary vocational education and economic development in less developed regions: A case study of Southern Xinjiang. China Vocational and Technical Education (24), 36-42.

[18] Jiahui Jin. (2020). Exploration on Diversification Path of Higher Education Investment. Academic Journal of Humanities & Social Sciences (9. 0). The doi: 10. 25236, AJHSS. 2020. 030905.