Effect of Introduction of Compulsory Education Law on the Rate of Return to Education

Haobin Wang[1, a, *], Zhuoxuan Liu[1, b], Tingting Gan[2, c]

[1]Central University of Finance and Economics, Beijing, 100081, China

[2]Nanchang Hangkong University, Nanchang, 330063, China

[a,*] 312736510@qq.com, [b]13261610866@163.com,[c] 1224038653@qq.com

Abstract. based on the data from China general social survey (cgss). Using method of DID and regression discontinuity to compute the influence on the returns to education due to Law on Compulsory Education. The results are various and the OLS regression simply indicates that the rate of return to education in china is only 5.24%. The DID model estimate that one more year of study in school will increase salary by 6.8%.However,the result of Regression Discontinuity design is that the rate of return to education surprisingly reaches to 46%.

Keywords: development economics; return to education (RTE); Regression Discontinuity; DID.

1. Introduction

Education is one of fundamental development objects which is the base towards wonderful life. Meanwhile education play a key role in economic development. In 1960s, Schultz proposed human capital theory that believed human capital is the most important resources. Investment in education is the major part in human capital investment. And the rate of return to education is generally thought to be the increase of personal income for one more year of education.

2. Prior Literature

A lot of researches have been published paper about return to education (RTE) and have been attempt to apply distinct ways and switch the focuses. The Mincer equation is the mainstream which many researchers expand equation to estimate the effect of policies such as gender, educational attainment, areas difference, family income etc. Chengjin Li and Yongzhen Zhang(2018)add income factor into the research and find out that at each education level the rich family get higher rate of RTE [1]. A creative method is that comparing the RTE between twins which could reduce the internal problem and get more precise estimations. The author draw a conclusion that traditional estimation is overestimated by 8 percent. [2]

Regression Discontinuity is also applied generally in studying the effect of policy. [3]Xuezheng Qin(2017) estimated the effect of one-child policy in china on individual future achievement and got the result of the policy positively influence one's career which indicate that families face the trade off between the quality and quantity of kids. [4]What's more, David Figlio(2018) uses a sharp RD design to estimate the effects of lengthening the school day for low-performing schools in Florida by exploiting an administrative cutoff for eligibility. Results indicate significant positive effects of additional literacy instruction on student reading achievement. In particular, we find effects of 0.05 standard deviations of improvement in reading test scores. [5]

3. Data

All data in this paper from China general social survey (cgss). Considering the Regression Discontinuity need relatively sufficient samples, samples from two year—2013 and 2015 are included in regression.

Deleting the invalid samples and transform the people whose income are zero to be one (in order to keep the distribution of the sample), the number of observations is 13507. Based on the law of compulsory education enacted in 1986, kids who are six or elder are qualified to enroll to the school.

So the cutoff which means kids effected by policy is set to be 1980.(the following part will illustrate more details about this).

4. Empirical Strategy

4.1 OLS Estimation with Crosssection Data

Based on the mincer equation:

$$\ln(\text{wage}_i) = \beta_0 + \beta_1\text{edu}_i + \beta_2\text{exp}_i + \beta_3\text{exp2}_i + \mu_i \quad (1)$$

$$\ln(\text{wage}_i) = \beta_0 + \beta_1\text{edu}_i + \beta_2\text{exp}_i + \beta_3\text{exp2}_i + \beta_4 X_i + \beta_5 X_i * \text{edu}_i + \mu_i \quad (2)$$

The wage represents annual income of a person.; edu is the year of education; exp is person's working experience measured by year since he/she graduated ; exp2 is quadratic of exp; Xi includes all other potential variables such as gender, health etc.

4.2 DID Estimation

DID is widely applied to evaluate the effect of policies or laws. By combining the time difference with the systematic difference between the experimental control group, this kind of "natural experiment " can add more external variables to estimate effect.

The core of DID is to calculate the double-difference estimator by comparing before and after (time) and cross section (control group and experimental group).

$$\text{d} = (Y_{Treatment,\text{t1}} - Y_{Control,t1}) - (Y_{Treatment,\text{t0}} - Y_{Control,t0})$$

d is the double-difference estimator which is the also the estimated policy effect. The foot marks treatment and control represent the experimental group and the control group respectively, while t0 and t1 represent before and after the release of the policy. In this paper, it is shown that the kids was born around 1986.

So the regression equation is that:

$$\ln(\text{wage}_i) = \beta_0 + \beta_1\text{T}_i + \beta_2\text{year}_i + \beta_3\text{X}_i + \mu_i \quad (3)$$

The T is a dummy variable which classify the sample based on whether the kid finished compulsory education.(T=1 also means the kid is belong to experiment group)

$$\begin{cases} T = 1 & graduated\ from\ junior\ high\ school \\ T = 0 & not\ graduated\ from\ junior\ high\ school \end{cases}$$

The year is another dummy variable to judge whether kids were born before the cutoff.

$$\begin{cases} year = 1 & \text{if birth} \geq 1980 \\ year = 0 & \text{if birth} < 1980 \end{cases}$$

X is the intersection of T and year, so the estimation of β_3 is the effect of compulsory education Law

$$\ln(\text{expense}_i) = \beta_0 + \beta_1\text{T}_i + \beta_2\text{year}_i + \beta_3\text{X}_i + \beta_4 edu_i + \beta_5 exp_i + \beta_6\text{exp2}_i + \mu_i \quad (4)$$

(4)is the equation after adding more control variables.

4.3 RD Estimation

Regression discontinuity was first proposed by Thistle Waite and Campbell in 1960 to estimate the treatment effect, which is similar to random experiment Considering the realistic possibilities, some districts after the implementation of the policy may not applied policy and some children may not go to school for various reasons, it implies that just probability has been raised with some level which could be found that treatment effect can be represented by the variation of year of schooling. So we adopt the method of fuzzy RD regression. 2SLS is standard way for fuzzy RD regression. Specifically, the first-stage and reduced form model are shown as follows:

$$\mathrm{edu}_i = \pi_0 + \pi_1 \mathrm{d}_i + h(D_i) + \mu_i \quad (5)$$

$$\ln(\mathrm{wage}_i) = \delta_0 + \delta_1 \mathrm{d}_i + g(D_i) + \gamma_i \quad (6)$$

The structure equation is:

$$\ln(\mathrm{wage}_i) = \beta_0 + \beta_1 \mathrm{edu}_i + f(D_i) + \varepsilon_i \quad (7)$$

h (), g() and f() are polynomial models of the driving variable birth (birth date of sample individuals) on both sides of the breakpoint.

$$\mathrm{Di} = \begin{cases} 1 \text{ if birth} \geq 1980 \\ 0 \text{ if birth} < 1980 \end{cases}$$

In the fuzzy RD regression model, in order to obtain the impact of policies, we must "widen" the gap by means of the average treatment effect:

$$\check{LATE} \equiv E[(y_1 - y_0) \mid x = c] = \frac{\lim_{x \downarrow c} E(y \mid x) - \lim_{x \uparrow c} E(y \mid x)}{\lim_{x \downarrow c} E(D \mid x) - \lim_{x \uparrow c} E(D \mid x)}$$

Which is the exactly the IV estimator in equation (5): $\beta_1 = \frac{\pi_1}{\delta_1}$

Whether it is affected by the universal obligation or not is only a strategic policy made by the government based on the development goals, and has nothing to do with other factors of ability that can affect the income. Therefore, we believe that D is an appropriate instrumental variable for the years of education.

5. Results

5.1 Descriptive Statistics

Table 1. descriptive statistic

Variable	Obs	Mean	Std. Dev.	Min	Max
gender	13507.000	0.534	0.499	0.000	1.000
birth	13507.000	1968.518	15.411	1920.000	1997.000
ethnic	13507.000	0.043	0.202	0.000	1.000
educ	13507.000	10.722	3.176	6.000	19.000
grad	13507.000	1986.851	16.474	1933.000	2014.000
wage	13507.000	374000.000	1830000.000	1.000	10000000.000
health	13507.000	0.888	0.316	0.000	1.000
exp	13507.000	27.149	16.474	0.000	81.000
T	13507.000	0.858	0.349	0.000	1.000
year	13507.000	0.280	0.449	0.000	1.000
X	13507.000	0.269	0.443	0.000	1.000
D	13507.000	0.280	0.449	0.000	1.000
lnwage	13507.000	7.303	4.918	0.000	16.118
exp2	13507.000	1008.428	998.133	0.000	6561.000

5.2 OLS

Table 2. OLS estimation

VARIABLES	OLS
educ	0.0524***
	(0.0146)
exp	0.120***
	(0.00851)
exp2	-0.00375***
	(0.000134)

The result of the mixed OLS equation(Mincer equation) is calculated return on education is 5.24%. That is, for every additional year of education, an individual's occupational income increased by 5.24%.

5.3 DID Estimation

Table 3. DID Estimation

VARIABLES	DID	DID	DID	DID
T	1.116***	-1.446***	-1.357***	-1.315***
	(0.128)	(0.144)	(0.145)	(0.145)
year	1.068***	-2.872***	-2.879***	-2.866***
	(0.412)	(0.381)	(0.385)	(0.383)
X	0.0684	1.880***	1.887***	1.868***
	(0.423)	(0.383)	(0.387)	(0.386)
educ		0.0760***	-0.00201	-0.123***
		(0.0178)	(0.0185)	(0.0222)
exp		0.0742***	0.0687***	0.0723***
		(0.0126)	(0.0128)	(0.0127)
exp2		-0.00344***	-0.00340***	-0.00338***
		(0.000170)	(0.000172)	(0.000171)
gender		2.046***		
		(0.0752)		
ethnic		0.175		
		(0.185)		
health		1.104***		
		(0.122)		
geneduc			0.156***	0.156***
			(0.00679)	(0.00677)
ethedu				0.0128
				(0.0170)
healedu				0.123***
				(0.0125)
Constant	6.029***	7.402***	9.452***	9.395***
	(0.116)	(0.344)	(0.328)	(0.327)
R-squared	0.020	0.224	0.208	0.213

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Equation (1) is the simplest DID model. The results show that due to the promulgations of the compulsory education law, individual labor income will increase by 6.8%, and significantly at the 1% level. Equations (2), (3) and (4) are the extension of the theory of mincer's equation. From the results in table 2, we can see that the RTE of female is almost zero, while the return on men's education is higher than that of OLS, reaching 15.6%.

5.4 RD Estimation

5.4.1 Average Treatment Effect

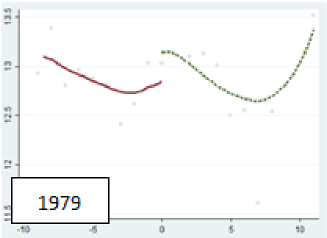

Fig 1. Cut off: 1979

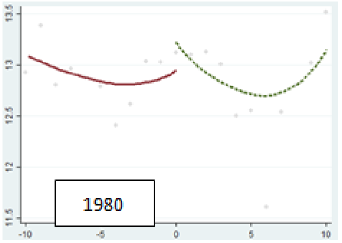

Fig 2. Cut off: 1980

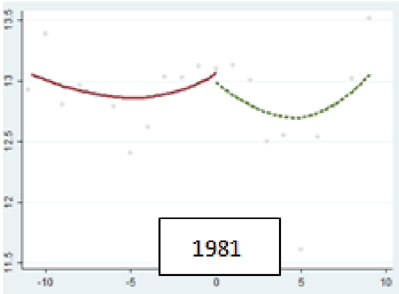

Fig 3. Cut off: 1981

The ordinate represents the average years of education and the year of birth, respectively. We can see that the average years of education of an individual and two years 1979 and 1980 (kids are 6 years old in 1986) are the focal points of school attendance for school-age children. Presents an inverted U shape, first decreases and then increases before the breakpoint, then there is a significant increase at the breakpoint, and then it still decreases and then increases. The average policy effect is to increase the annual occupational income by 44,300 Yuan.

5.4.2 2SLS Estimation

Table 4. RD Estimation

	(1)	(2)	(3)
VARIABLES	first-stage	2sls	2sls
D	-2.750***		
	(0.0845)		
exp	-0.353***	0.209***	0.207***
	(0.00645)	(0.0124)	(0.0128)
exp2	0.00344***	-0.00456***	-0.00454***
	(9.04e-05)	(0.000164)	(0.000165)
health	0.404***	0.883***	
	(0.0716)	(0.138)	
gender	0.339***	1.860***	
	(0.0440)	(0.0800)	
ethnic	-0.380***	0.363*	
	(0.108)	(0.192)	
educ		0.467***	0.438***
		(0.0452)	(0.0507)
geneduc			0.0872***
			(0.00991)
Constant	17.09***	-0.568	1.051
	(0.137)	(0.649)	(0.691)
R-squared	0.361	0.156	0.135

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

To our surprise, universal compulsory education reduces the length of schooling of individuals by 2.75 years and passes the significance test at 1% level. Equation (2) shows that the compulsory education law enacted at the breakpoint has a 46% increase in the individual return on education, and the income of men is more affected than that of women, while the healthy people are more affected. In (3) the gender differences implies the return on education of men increased by 8.72% more than that of women.

6. Conclusion

Based on the results of simple OLS regression, the classical mincer equation and its extension equation, the RTE in China is estimated to be 5.24%.

The results of the DID model show that the income of individual will increase by 6.8% due to the policy which is greater than OLS and t shows the positive effect of law on RTE. The difference is that the return on education for women is almost zero, while that for men is higher than OLS, reaching 15.6%. Gender differences are the same in the OLS and DID models and are significant both gender differences in educational returns are 15.6%.

The results in the RD show that the effect of the enactment of law on the individual's RTE is much greater than previous two models. From the perspective of average effect, there was a significant increase at the breakpoint, increasing the annual occupational income by 44,300 yuan. The results of 2sls regression show that compulsory education reduces the number of years of education by 2.75 years .Enacting compulsory education at the cutoff increased individual educational returns by 46%, and it affected men more than women in terms of income, and the healthy is more suffered. Comparing the differences in educational returns between men and women. As a result of the policy, the return on education of men increased by 8.72% more than that of women.

In the empirical results of the three models, we find that although there are differences in the calculation of educational return, the gender difference of educational return is relatively stable, reaching about 10%. Contrary to popular belief, men's returns on education are higher than women's.

The innovation of this paper is the calculation and comparison of different educational returns between different models and the focus on gender differences and the good regression results obtained, as well as the use of RD model to calculate the effect of policy effect. After adding new explanatory variables and interaction terms in the robustness test, it actually does not reach an exaggerated 46.7% but 36.3%.

Acknowledgements

The authors gratefully acknowledge the support from China general social survey (cgss) for the data provided.

References

[1]. ChengJin Li, YongZhen Zhang. Comparative study on educational return rate of poor and non-poor families -- based on a survey of 1,749 farmers in 14 poverty-stricken villages in gansu province [J]. Northwest Population on Journal, 2008,39 (03) : 109-117.

[2]. ZhiJun Sun. Estimation of educational return based on twin data [J]. China Economic Quarterly, 2014,13 (03) : 1001-1020.

[3]. JingZen Yu, Chunchao Wang. The rise of new quasi-random experiment method: breakpoint regression and its application in economics [J]. Economic Perspectives, 2011 (02) : 125-131.

[4]. Xuezheng Qin. Does the one-child policy improve children’s human capital inurban China? A regression discontinuity design[J]. Journal of Comparative Economics ,45 ,(2017), 287–303.

[5]. David Figlio. Do students benefit from longer school days? Regression discontinuity evidence from Florida's additional hour of literacy instruction [J] Economics of Education Review, 67, (2018),171-183.