Estimates of the Returns to Schooling in Taiwan: Evidence from a Regression Discontinuity Design

Junchao Zhang*

January 11, 2019

Abstract

Using micro-data from the Manpower Utilization Survey 1995-2015 of Taiwan, this paper exploits exogenous variation in years of schooling caused by the 9-year compulsory education reform to estimate the economic returns to schooling. The results of fuzzy regression discontinuity design indicate that the causal effect of education on wage is approximately 5.6%, smaller than the corresponding ordinary least squares estimates. Our results are robust to various specifications and placebo tests.

JEL classification: I26;I28

KEY WORDS: returns to education, regression discontinuity, Taiwan

* *The Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa, Tokyo 190-8562, Japan E-mail address: ztempest0218@gmail.com*

1 Introduction

Education has a crucial implication in human society, as it has long been of interest to policymakers, economists, and individuals themselves. From a national perspective, policymakers often assert that public expenditure on education may lead to poverty reduction, and promote economic growth (Hanushek and Woessmann, 2008). From an individual perspective, students and their families also make educational choices with financial resources that they are able or willing to afford. With a limited budget, parents tend to invest more in the children who are expected to have higher earnings as adults.

Despite the rapid accumulation of literature on the return to education in Taiwan, little work has been done to investigate causality. Following Mincer (1974), Wu (2003) estimates the Mincer wage equation and have found the return to education by 5.0% - 7.8% in Taiwan during 1978 to 2001. Chuang and Chao (2001) find that returns to schooling are heterogeneous across samples with different final educational attainments[1], and females have higher returns to schooling than males in Taiwan.

However, the causality between education and economic outcomes is complicated by endogenous educational decisions. If unobservable factors are common between schooling and wages, then the estimated education return is biased. Since Angrist and Krueger (1991), economists have devoted a great deal of attention to correcting the omitted variable (ability) bias by instrumental variable (IV) method (Harmon and Walker, 1995; Patrinos and Sakellariou, 2005; Pischke and Wachter, 2008; Oreopoulos, 2007) or fuzzy regression discontinuity design (Devereux and Hart, 2010).

Using a large Taiwanese data set, we establish the causality between schooling and wages by a fuzzy regression discontinuity design and our results are different from previous IV studies. Specifically, Taiwan carried out a compulsory education reform in 1968, the mandatory years of schooling were extended to 9 years from 6 years. Due to the compulsory schooling law, those who were born before and after 1955 have different probabilities to leave schools, thus have different years of schooling (Zhang, 2018). Exploiting this natural experiment, we show a very moderate return to education by approximately 5.6%, while previous IV studies show returns to education by 10% - 15%

[1] 2.30% for senior high school, 3.98% for vocational school, 4.58% for junior college, and 12.20% for university.

(Card, 2001)[2].

The remainder of this paper is organized as follows. In Section 2, we present the identification strategy used to estimate returns to education. Section 3 describes the data and descriptive statistics. Section 4 discusses the results of OLS and IV. Section 5 provides robustness checks for our results. Section 6 concludes.

2 Identification Strategy

To address the endogeneity of educational choice, we use the change in compulsory schooling law to identify the causal effect of schooling on wages by a fuzzy regression discontinuity design. The new policy that raised mandatory years of schooling from 6 to 9 in 1968 generates a discontinuous change in individuals' educational outcomes across cohorts born before and after 1955. Using this exogenous variation in schooling, the estimate of interest, δ_1, can be defined in terms of limits under the local average treatment effect (LATE) framework as follows (Imbens and Angrist, 1994):

$$\delta_1 = \lim_{\Delta \to 0} \frac{E[w_{it}|X_0 < X_{it} < X_0 + \Delta, \boldsymbol{Z}_{it}] - E[w_{it}|X_0 - \Delta < X_{it} < X_0, \boldsymbol{Z}_{it}]}{E[S_{it}|X_0 < X_{it} < X_0 + \Delta, \boldsymbol{Z}_{it}] - E[S_{it}|X_0 - \Delta < X_{it} < X_0, \boldsymbol{Z}_{it}]} \tag{1}$$

where w_{it} and S_{it} represent log monthly wage and schooling, respectively. X_{it} is the year of birth, and X_0 is 1955, which is the first cohort will be affected by the new compulsory schooling law. $\boldsymbol{Z}_{it}$ is a vector of control variables including gender, marital status, year dummies, tenure, and tenure squared. Equation (1) captures the causal effect on compliers, defined as individuals whose eligibility for the new policy changes as we just move the value of X_{it} from $X_{it} - \Delta$ to $X_{it} + \Delta$, where Δ is very close to X_0. As Imbens and Lemieux (2008) suggest that using a high-order polynomial function could lead to misleading results in regression discontinuity design, we use the local linear regression approach to estimate δ_1 alternatively. In practice, a two-stage least squares (2SLS) system of equations is used to perform the fuzzy regression discontinuity design (Hahn et al., 2001).

[2] Card (2001) offers a survey of IV literature on returns to education, the IV estimates of returns to education are typically larger than the corresponding OLS estimates.

3 Data

The data used in this research is Manpower Utilization Survey 1995-2015 of Taiwan. The survey has been collected annually by a stratified two-stage random sampling procedure since 1979, and each wave is a cross-sectional data. The survey contains basic information on individual characteristics and labor market activities. This research uses the pooled cross-sectional data of 20 years between 1995 and 2015 for following reasons. First, educational attainment over university level can not be distinguished before 1995. Second, we have to ensure that the observed educational attainment is not censored at a survey year. After 1995, the initially eligible cohorts of the 9-year compulsory schooling law would be at least 35 years old, and have few probabilities to attend school.

Each wave of Manpower Utilization Survey contains approximately 50,000-60,000 individuals. For the analysis of returns to education, we restrict the sample as follows: (1) Monthly wages are used to measure labor incomes, observations with missing wages are dropped. (2) We only use observations whose information on year-of-birth, gender, marital status, educational attainment is not missing. Year-of-birth is the forcing variable that exogenously determines schooling across cohorts. (3) As tenure could be an important input of wages, we only use individuals whose tenures are not missing. The data offer exact tenure variable with the accuracy of months. We convert the tenure variable into decimal years. After the above procedures are finished, we get a final sample with 601,689 workers[3].

-Insert Table 1-

The descriptive statistics of the final samples used to estimate the OLS and fuzzy RD are shown in Table 1. Full sample is used to estimate the OLS and global polynomial RD, discontinuity samples with 3-year, 4-year and 5-year bandwidths are used to estimate the local linear RD. The discontinuity samples are very similar across different bandwidths. From Column (3), the average age of the discontinuity sample is approximately 48.3, 50% of them benefit from the 9-year compulsory schooling law, and the average tenure is 12.3 years . As the discontinuity samples contain those born just before and after 1955, approximately 84% of them are married in the survey years. About 67% are males in the discontinuity samples because males have higher labor force participation rate than

[3]No observations are dropped through procedures (2) and (3).

females in Taiwan.

4 Results

In this section, we discuss the estimates of the fuzzy regression discontinuity design, which were designed analyze returns to education in Taiwan. We first discuss the first stage results, and then compare the OLS and IV estimates. From top to bottom in Table 2, three panels are listed for specifications with different covariates. We report White-Hubser standard errors due to the heteroskedasticity.

4.1 First Stage

Figure 1 provides a graphical evidence of the first stage, which shows the effect of the compulsory schooling law on years of schooling. The figure shows a clear discontinuous change in average schooling around the eligibility cutoff. Optically, schooling is 0.3-0.4 years longer among initially eligible cohorts. While we can only measure the eligibility at the year-of-birth level, we also use 1980 Population and Housing Census of Taiwan to examine the jump at the cutoff point. Out first stage results from Manpower Utilization Survey (Figure 1) are consistent with the results from the census (Figure A1), showing a discontinuity at 1955. According to local linear regression estimates in Table 2, schooling are longer among initially cohorts by 0.17-0.32 years (Columns 5, 7, and 9).

-Insert Figure 1-

4.2 IV and OLS estimates

From Columns (2), (4), (6) and (8) of Table 2, we find a significant positive education return for IV estimates, regardless of functional forms or different local linear bandwidths used. From Panel C in Column (6), education is estimated to raise the wage by approximately 5.6%, when tenure variables (tenure and tenure squared) are controlled for in regression. However, due to some data limitation, few IV studies of the returns to education has controlled for tenure. Panel A and Panel B in the same column show that an additional year of schooling is estimated to increase the wage by only approximately 4.2% when tenure variables are not controlled for. Similar patterns can also be observed in other specifications of fuzzy regression discontinuity design in columns (2), (4) and (8).

-Insert Table 2-

While our OLS estimates are comparable to the literature, our IV estimates only indicate very moderate returns to education. This magnitude of our IV estimates relative to the OLS is different with most previous studies. Specifically, our OLS estimates consistently suggest return to education by 5.6-6.2% across different specifications in Column (1), very close to the 5.2% returns to education of 1940-49 cohort for U.S. males (Angrist and Krueger, 1991) and the 6.1% returns to education for UK males. (Harmon and Walker, 1995). On the contrary, our IV estimates are only 5.6% from Panel C in Column (6), smaller than those in U.S. (7.8% for 1940-1949 cohort) and in UK (15.3%). We also construct placebo reforms to examine the robustness of the fuzzy RD estimates. Table 3 shows that placebo reforms of 1952, 1954, 1956 and 1958 have no effect on the wages across different bandwidths. The eligibility is not related to gender, marital status, and tenure (Table A1).

-Insert Table 3-

5 Concluding Remarks

In this study, we exploit the discontinuity in years of schooling induced by an extension of compulsory schooling law to estimate the returns to education. We find that the eligibility of 9-year compulsory education increases the average schooling by 0.26, among those who acquired additional schooling as a result of the reform. Subsequently, we show that this discontinuity in years of schooling leads to a significant increase in wages. An additional schooling is estimated to raise the wage by approximately 5.6%. Contrast to previous IV studies, the fuzzy RD estimates (essentially an IV method) in this research are smaller than the corresponding OLS estimates. Our results are robust to different specifications and placebo tests.

References

Angrist, J. D. and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics*, 106(4):979–1014.

Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems. *Econometrica*, 69(5):1127–1160.

Chuang, Y.-c. and Chao, C.-Y. (2001). Educational choice, wage determination, and rates of return to education in taiwan. *International Advances in Economic Research*, 7(4):479–504.

Devereux, P. J. and Hart, R. A. (2010). Forced to be rich? returns to compulsory schooling in britain. *The Economic Journal*, 120(549):1345–1364.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica*, 69(1):201–209.

Hanushek, E. A. and Woessmann, L. (2008). The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*, 46(3):607–668.

Harmon, C. and Walker, I. (1995). Estimates of the economic return to schooling for the united kingdom. *American Economic Review*, 85(5):1278–1286.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Imbens, G. W. and Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615 – 635.

Mincer, J. (1974). *Schooling, Experience, and Earnings.* National Bureau of Economic Research, Inc.

Oreopoulos, P. (2007). Do dropouts drop out too soon? wealth, health and happiness from compulsory schooling. *Journal of Public Economics*, 91(11-12):2213 – 2229.

Patrinos, H. and Sakellariou, C. (2005). Schooling and labor market impacts of a natural policy experiment. *LABOUR*, 19(4):705–719.

Pischke, J.-S. and Wachter, T. v. (2008). Zero returns to compulsory schooling in germany: Evidence and interpretation. *The Review of Economics and Statistics*, 90(3):592–598.

Wu, H. (2003). Returns to schooling in taiwan, 1978-2001. *Taiwan Economic Forecast and Policy*, 33(2):97 – 130.

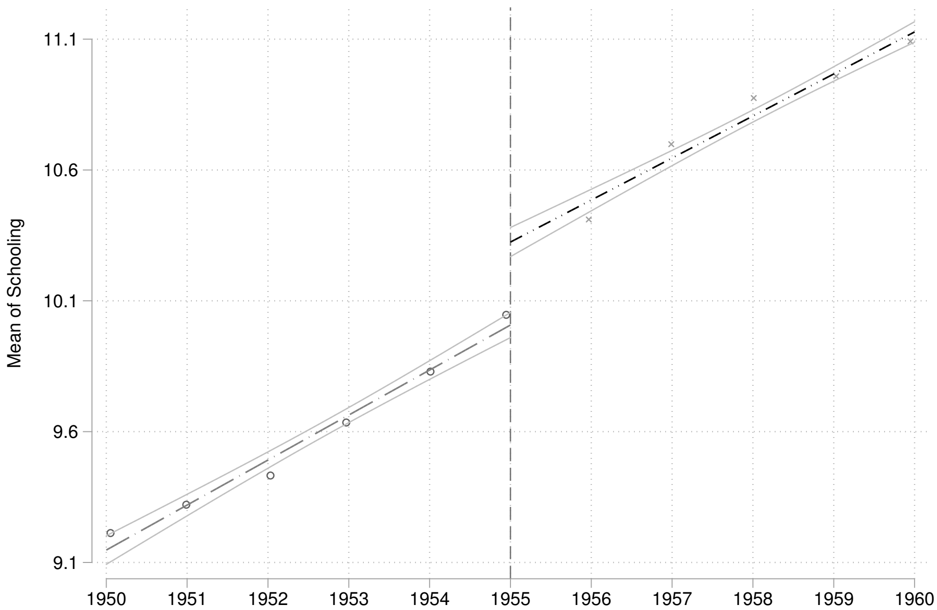
Zhang, J. (2018). The impact of 9-year compulsory education: quasi-experimental evidence from taiwan. *Applied Economics*, 50(45):4866–4878.

Table 1: Descriptive Statistics

VARIABLES	(1) Full Sample	(2) ±3 Sample	(3) ±4 Sample	(4) ±5 Sample
Year of birth	1965.04	1955.18	1955.34	1955.55
	(12.74)	(1.99)	(2.55)	(3.08)
Years of schooling	11.55	10.26	10.29	10.33
	(3.60)	(3.63)	(3.61)	(3.60)
log wage	10.30	10.39	10.38	10.38
	(0.55)	(0.59)	(0.59)	(0.59)
Eligible (=1 if born after 1955)	0.77	0.47	0.50	0.53
	(0.42)	(0.50)	(0.50)	(0.50)
Age	39.74	48.42	48.28	48.07
	(11.97)	(6.11)	(6.27)	(6.45)
Male	0.62	0.67	0.67	0.67
	(0.49)	(0.47)	(0.47)	(0.47)
Married	0.63	0.85	0.84	0.84
	(0.48)	(0.36)	(0.36)	(0.36)
Tenure	8.84	12.35	12.28	12.20
	(9.15)	(8.99)	(9.00)	(8.99)
Observations	601,689	94,052	120,616	145,952

Data source: Manpower Utilization Survey 1995-2015 of Taiwan

Figure 1: Average Years of Schooling by Year-of-Birth



Notes: Author's calculation by Manpower Utilization Survey 1995-2015 of Taiwan

Table 2: Fuzzy RD Estimates of Education Return

		Polynomial		Local Linear					
		Quadratic		Discontinuity ± 3		Discontinuity ± 4		Discontinuity ± 5	
	OLS	IV	1st Stage	IV	1st Stage	IV	1st Stage	IV	1st Stage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES	log wage	log wage	Schooling	log wage	Schooling	log wage	Schooling	log wage	Schooling
Panel A: Basic controls									
Years of schooling	0.067***	0.060***		0.040***		0.042***		0.041***	
	(0.000)	(0.001)		(0.003)		(0.003)		(0.002)	
Eligible			2.027***		0.177***		0.264***		0.322***
			(0.016)		(0.051)		(0.043)		(0.038)
Panel B: Basic controls+Year FE									
Years of schooling	0.067***	0.056***		0.040***		0.042***		0.041***	
	(0.000)	(0.001)		(0.003)		(0.003)		(0.002)	
Eligible			0.616***		0.177***		0.264***		0.321***
			(0.018)		(0.051)		(0.043)		(0.038)
Panel C: Basic controls+Year FE+Tenure									
Years of schooling	0.068***	0.062***		0.054***		0.056***		0.056***	
	(0.000)	(0.001)		(0.003)		(0.002)		(0.002)	
Eligible			0.637***		0.175***		0.261***		0.317***
			(0.018)		(0.051)		(0.042)		(0.038)
Observations	601,689	601,689	601,689	94,052	94,052	120,616	120,616	145,952	145,952

Notes: Standard errors in parentheses. In Panel A, basic demographic variable of gender and marital status are controlled. In Panel B, basic controls and year fixed effect are included. In Panel C, tenure and tenure squared variables are additionally controlled. *** p<0.01, ** p<0.05, * p<0.1

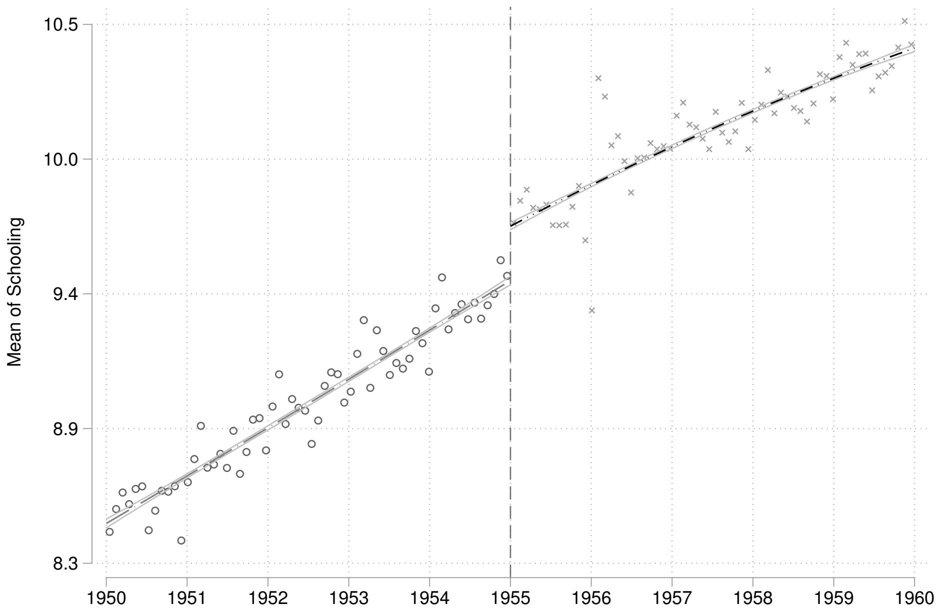
Table 3: Reduced Form Estimate of Placebo Treatment on Wage

VARIABLES	±3 Sample (1) log wage	±4 Sample (2) log wage	±5 Sample (3) log wage
Panel A: -3 Years (1952)			
Placebo Treatment	0.010	0.005	0.005
	(0.009)	(0.008)	(0.007)
Observations	77,164	98,091	118,168
Panel B: -1 Years (1954)			
Placebo Treatment	-0.010	-0.001	-0.000
	(0.008)	(0.007)	(0.006)
Observations	89,318	113,044	137,535
Panel C: +1 Years (1956)			
Placebo Treatment	0.006	0.004	-0.003
	(0.008)	(0.006)	(0.006)
Observations	98,488	126,960	154,032
Panel D: +3 Years (1958)			
Placebo Treatment	-0.002	0.001	-0.004
	(0.007)	(0.006)	(0.005)
Observations	107,591	137,515	166,989

Note: Robust standard errors in parentheses. Gender, marital status, year fixed effect, tenure, tenure squared are included in all regressions. Each coefficient denotes a separate regression. *** p<0.01, ** p<0.05, * p<0.1

Appendix A

Figure A1: First Stage Evidence from Census



Notes: Author's calculation by 1980 Population and Housing Census of Taiwan.

Table A1: Correlation Between Eligibility and Covariates

	Discontinuity ± 4		
	(1)	(2)	(3)
VARIABLES	Eligible	Eligible	Eligible
Year of birth	0.170***	0.170***	0.170***
	(0.000)	(0.000)	(0.000)
Age	0.000	0.000	0.000*
	(0.000)	(0.000)	(0.000)
Male	-0.001	-0.001	-0.001
	(0.002)	(0.002)	(0.002)
Married	0.002	0.002	0.002
	(0.002)	(0.002)	(0.002)
Tenure			0.000
			(0.000)
Tenure squared			-0.000
			(0.000)
Constant	-331.504***	-332.179***	-332.182***
	(0.417)	(0.601)	(0.601)
Year FE	No	Yes	Yes
Observations	120,616	120,616	120,616
R-squared	0.751	0.751	0.751

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1