Analysis Of Chinese National College Entrance Examination's Effect on Labor Outcomes and Happiness

Yunshu Shao *

Department of Economics, Korea University, Seoul, Korea

* Corresponding Author Email: cristalshao521@korea.ac.kr

Abstract. This study evaluates the impact of earning an undergraduate degree on one's level of happiness, work prospects, income, and subjective life quality based on the logical outcome that one's college entrance exam score exceeds the threshold for undergraduate admission. This study's primary empirical methodology is the Regression Discontinuity Design (RDD), which examines undergraduates of science and art fields individually and urban and rural subgroups in terms of their return to Chinese higher education. Under the Chinese College Admission System (CAS), the National College Entrance Examination (NCEE) scores are the most essential benchmark for enrolling students, which are the well-defined cutoffs for applying the RDD method. More specifically, the study focuses on two aspects of "a better life": material and mental. The material aspect is indicated by the yearly income and the probability of getting a job. The level of happiness and subjective life quality reflects the mental aspect. The results show that having an undergraduate degree or passing the score line does not significantly impact life quality and labor market outcomes. However, the result still indicates that higher scores can make people earn more and happier. Meanwhile, age seems to be a more critical factor in explaining the labor market outcomes. At the same time, life quality is likely to be determined by factors other than academic performance and working experience.

Keywords: National College Entrance Examination; Regression Discontinuity Design.

1. Introduction

1.1. Background

During the past 30 years, the Chinese economy has maintained a high growth rate. The number of people enrolling in higher education climbed from 0.27 million to 1.08 million between 1977 and 1998, growing at a 6.8% annual pace. The annual growth rate of university admissions later increased to 12.4% between 1998 and 2015. On the one hand, the economy's quick growth increased the need for workers with advanced degrees. On the other hand, the market for higher-educated workers has grown significantly due to the expansion of higher education, and this rise in supply may have distinct potential benefits for those who choose to pursue further education or not. The value of an undergraduate degree in terms of return on investment is one of the questions to be answered. Additionally, since education is meant to impact both job and life, the return should be split into these two categories. It is also important to consider whether rates of return differ between urban and rural areas and between science and arts subjects.

1.2. Related Research

So far, many researchers have studied the relevant questions via different methods. For example, by the OLS method, there is a study that estimates the rate of return to studying subjects with different levels for first-year undergraduate, masters, and PhDs in Britain, having found a "quality effect" that more capable students will self-select into more challenging subjects and can earn more money than others [1]. Then, another focuses on the return to education and its distribution in urban China. Except for using quantile regression (QR) and having three types of control variables: parents' background, industry, and education level, the study has also compared the results from QR with those of OLS methods [2]. In 2010, Fan et al. studied the causal effect of university education on earnings by Regression Discontinuity Design (RDD) method, finding that being eligible for university admission implies a significant increase in earnings and four-year university students generally earn more than those with three-year college or high school degree [3]. Considering the long-term change in returns to higher education in China, the researcher has created a Sample Selection Model employing both the IV method and the Heckman two-stage procedure to address potential endogeneity bias and sample selection bias and then found that results are higher than OLS estimates but lower than IV estimators of many previous literature [4].

In terms of methods, some researchers also point out potential problems of specific models, trying to make the estimations unbiased. One early research that presents IV estimation of the returns to schooling for Honduran males has pointed out that the OLS model can suffer from bias since it does not take individual heterogeneity and the endogeneity of schooling investment into account [5]. That is, educational outcomes are not randomly assigned throughout the population, thus methods other than OLS should be applied to tackle the self-selection and endogeneity issue. One typical self-selection issue is that, if more capable students are attracted to higher education, then part of the return to higher education may reflect a quality effect. In the meanwhile, some study has also used the Quantile Regression to control endogeneity and heteroscedasticity.

Compared with other methods, RDD makes it possible to estimate the average treatment effect in environments in which randomization is unfeasible and control the estimation bias caused by endogenous problems. As for this methodology, the very first study using RDD to estimate the causal effect of university education on earnings is done by Fan et al. (2010), which applies fuzzy RDD method. Also, by using RDD method, researchers can study the jump (change of intercept) and kink (change of slope) around the cutoff. For example, a paper using RDD method to evaluate the effects of two social programs on education, which are elite school attendance in the UK and compulsory military service in Russia. In both cases, the program's causal effect on education is a kink rather than a jump [6]. Generally, RDD allows researchers to study the effect of passing a certain cutoff.

The studies cited above have used plenty of methods to study educational issues. Many studies tend to address the biases caused by OLS model, finding ways to avoid such problems caused by heteroskedasticity and endogeneity. Among all the methods, RDD seems to be a more plausible method in research related to NCEE since 1) passing the score line or not is an important indication of academic performance, which is a feasible cutoff; 2) randomization is not highly possible when studying educational issues since endogeneity and heteroscedasticity are hard to avoid; 3) the regression results and graphs of RDD model can show both the effect of passing the cutoff and the distance from the cutoff. Since this study focuses on the college entrance exam in China, taking advantage of RDD method can discover the influence of test scores on labor outcomes and life quality.

However, sometimes the estimation of sharp RDD using score line as a cutoff can also be biased, considering passing the score line can affect multiple factors influencing future academic performance. An example is Loyalka et al. (2012), who compared senior university students whose college entrance exam scores were located on two sides of the score line of first-tier universities in China. They found that students just above the score line were less likely to be admitted to their ideal major than those just below the score line. They mention that implies that the estimated effect of passing the score line on earnings would be biased, because studying a major that students do not prefer may reduce their efforts in university, which in turn affect their future incomes [7]. This study indicates that sharp RDD seems to be limited since there might be unexpectable other effects around the cutoff. Fuzzy RDD with an IV estimator may solve this problem. To avoid the bias, this research will first use fuzzy RDD with passing the score line as the IV variable. Hence, the primary aim of this study becomes studying the effect of having an undergraduate degree on labor outcomes and life quality. Also, in the part of sharp RDD, this research uses the undergraduate score line, instead of the first-tier school score line.

Most previous studies focus on the effect of education on labor outcomes. Though a few studies analyze education's impact on factors related to life quality, they seldom study the effect of NCEE scores on life quality. For example, Behrman et al. (2015a) relied on within-monozygotic-twins fixed effects and showed that more years of schooling improved mental health and reduced smoking in

China [8]. Also, another study by Juncal Cuñado et al. means of estimating Ordinal Logit Models has shown that the educational level of individuals has a positive and significant effect on happiness [9]. Though the studies contribute to the research on the relationship between education and life quality, they lack the focus on the test score or passing the score line, which is crucial in the Chinese educational system. To fill the gap of previous studies, apart from labor outcomes, this study also considers the return on mental aspects such as happiness and subjective life quality. This research attempts to do so because, intuitively, the level of school a student attends to influences the level of education, which may contribute to life quality. Before analyzing labor outcomes and life quality, this study assumes that test scores positively affect labor outcomes and life quality to different extents.

When using RDD model, year of taking NCEE should be controlled since there are some policies related to the admission rate of universities. A typical example is the Chinese university admission expansion in the decade starting in 1999. A recent study by Huang et al. (2022) shows that the expansion improves individuals' earnings via increasing years of schooling [10]. Therefore, if the year of taking NCEE is not controlled, the estimation results will be biased.

1.3. Objective

This study extends the previous methodology by using Fuzzy RDD, with passing the undergraduate score-line as the instrumental variable (IV), considering the factor that passing the undergraduate score-line relates with, though it does not necessarily guarantee that an individual gets an undergraduate degree. For individuals passing the score line, having an undergraduate degree is supposed to affect labor outcomes and life quality since some occupations only employ people with undergraduate degrees or higher. Hence, this method is necessary to avoid a correlation between endogenous variables and the residual. By applying IV in this research, this research may discover something new about the relationship between passing the score line and labor outcomes/ life quality. Also, sharp RDD will later be used to see whether passing the score line will directly affect an individual's labor outcomes and life quality.

2. Method

2.1. Identification Methodology and Regression Discontinuity Design

The primary empirical strategy used in this paper is the Regression Discontinuity Design (RDD), studying an individual's return to Chinese higher education for urban and rural subgroups as well as for science and arts subject students. Under the Chinese College Admission System (CAS), the National College Entrance Examination (NCEE) scores are the most essential standard to enroll students, which means the score line of qualified universities can be well-defined cutoffs to apply RDD method. Throughout the study, the individuals who do not pass the cutoff is a valid counterfactual to individuals who pass the cutoff. It is also assumed that there is no full manipulation around the cutoff.

This study uses RDD, considering that the return on education is complicated and involves many endogeneities. The endogeneities are from at least three aspects. First, some congenital factors such as family background and personal gifts will make individuals over or underestimate the return of education. Also, as mentioned in a paper titled 'The Returns to College Education: A National Experiment from Expansion of Higher Education in China', education may have a signal effect, the effect that higher degree of education gradually becomes a threshold of entering big companies [6]. Such kind of signal effect may magnify the return on education itself. In this case, OLS is not suitable to solve the endogenous problem. However, RDD is not only possible to estimate the weighted average treatment effect in environments in which randomization is unfeasible but also control the estimation bias caused by endogenous problems.

Therefore, to address the endogeneities, which can be a problem due to the limitation of the widely used OLS method, this study uses a Fuzzy RDD method with 2SLS regressions. Using the dummy variable - passing the score line as an instrumental variable (IV), this study tries to focus on the relatively pure educational effect on the return of getting an undergraduate degree. However, sharp RDD is also used to more comprehensively understand the relationship between academic performance and the outcomes. Using a sharp RDD method, this paper is based on the identification methodology that people with college entrance examination scores just below the score line of undergraduate school represent a valid counterfactual for those with scores just above the score line of undergraduate school. The decision to use both fuzzy and Sharp RDD will be discussed in the "Result and Discussion" session.

In this case, a very important assumption is that the assignment near the cutoff point should be close to a randomized experiment. This assumption, which needs no 100% manipulation of the running variable, will be tested later through the validity test part.

In the Fuzzy RD design, there is a change in the probability of treatment passing the cutoff: $T^+ - T^-$ where T^+ and T^- are the treatment probabilities:

$$T^+ = \lim_{S \downarrow S_0} E[T_i|S_i = S_0] \text{ and } T^- = \lim_{S \uparrow S_0} E[T_i|S_i = S_0]$$

The running variable S is the individuals' actual college entrance examination scores, and the cutoff S_0 is the score line of undergraduate school. Then, the probability or intensity of treatment jumps at the cutoff is:

$$P[T_i = 1|S_i] = \begin{cases} g_1(S_i) \; if \; S_i \geq S_0 \\ g_0(S_i) \; if \; S_i < S_0 \end{cases}$$

where $g_1(S_i) \neq g_0(S_i)$, and the average treatment effect at the cutoff S_0 should be the differences between the limits of outcome (Y) on each side of the discontinuity recovered by the differences between the treatment probabilities:

$$ATE = \frac{Y^+ - Y^-}{T^+ - T^-}.$$

Thus, fuzzy RD coefficient has the interpretation of instrumental variable estimator: it is a ratio of reduced form coefficient to first stage coefficient. Besides, in this study, only the local average treatment effect (LATE) can be estimated, since compliers will get the treatment only if $S_i \geq S_0$, but they will not get the treatment if $S_i < S_0$.

2.2. Estimation Framework

With the second order polynomial in the parametric approach as the baseline specification, the basic estimation framework in this study is as follows:

$$Undergraduate_i = \begin{cases} 1, if \; the \; individual \; has \; an \; undergraduate \; degree \\ 0, \quad if \; the \; individual \; does \; not \; have \; an \; undergraduate \; degree \end{cases}$$

First Stage:

$$Undergraduate_i = \lambda + f(\mathbb{S}) + f(\mathbb{S}) \times D_i + \gamma D_i + X + v_i$$

Reduced Form:

$$Y_i = \gamma + \delta \cdot \widehat{Undergraduate_i} + f(\mathbb{S}) + f(\mathbb{S}) \times D_i + \rho D_i + X + u_i$$

where $f(\mathbb{S}) = \beta_1 \mathbb{S} + \beta_2 \mathbb{S}^2$, the trend relation of $E[Y_{0i}|\mathbb{S}]$.

In the case where sharp RDD is applied (discussed later), the framework is as below:

$$Y_i = \gamma + f(\mathbb{S}) + f(\mathbb{S}) \times D_i + \rho D_i + \alpha X + u_i$$

Y_i : outcome variables - income, probability of employment, the level of happiness, and subjective life quality

S_i: running variable - individuals' college entrance examination scores.

$\mathbb{S}$: normalized running variable, $\mathbb{S} = (S_i - u)/10$, u is the undergraduate score line.

D_i: dummy variable, indicating whether an individual passes the undergraduate score line.

$$D_i = \begin{cases} 1, & \mathbb{S} \geq 0 \\ 0, & \mathbb{S} < 0 \end{cases}$$

The Fuzzy RD coefficient δ is the main coefficient of interest: it measures the local average treatment effect of getting an undergraduate degree or not on the outcome variables. To get an unbiased estimator of the Fuzzy RD coefficient, the control variables are factors that can affect the normalized cutoff.

3. Data

As for the data of annual college entrance examination score lines (i.e., the cutoff) for each province, available data of score lines is collected from the published books, newspapers, and websites, such as web-pages of People Daily during the period (1977-2014) of exam used in the study, finding that the definition of undergraduate admission score line was not always consistent. Before 1992, there were only two admission score lines of universities in each province per year – one for the arts subject and the other for the science subject. The score lines can be directly used as the cutoff. After 1992, there are in general six different admission score lines for three levels: 1) the first-tier universities, 2) the second-tier universities, 3) the third-tier colleges or three-year colleges. For individuals taking exams after 1992, the second-tier universities' score lines are viewed as the cutoff because the third-tier colleges are more like technical colleges than universities.

Together with the data of NCEE scores lines for each province during 1977-2014, this study also uses the China Household Income Project Survey (CHIPs) data set in 2013. After primary data processing, the data has 4002 individuals, all of whom have taken NCEE in China. In terms of region, there are 1691 observations from the urban subgroup, and 2295 observations from the rural subgroup; the rest are working abroad or do not have a registered permanent residence. In terms of the subject of the exam, there are 2034 observations from science subject subgroup and 1968 observations from arts subject subgroup. The descriptive statistics of the whole sample and all subsamples are shown in Table 1.

Table 1. Descriptive Statistics

	Whole sample	**Urban residents**	**Rural residents**	**Science subject**	**Arts subject**
A. Outcome Variables:					
Undergraduate degree	0.405	0.511	0.327	0.453	0.355
	(0.491)	(0.500)	(0.469)	(0.498)	(0.479)
Employment probability	0.652	0.668	0.641	0.631	0.674
	(0.476)	(0.471)	(0.480)	(0.482)	(0.469)
Log(income)	10.342	10.493	10.221	10.391	10.294
	(0.793)	(0.758)	(0.802)	(0.794)	(0.790)
Level of happiness (normalized, between 0 and 1)	0.664	0.676	0.657	0.663	0.665
	(0.472)	(0.468)	(0.475)	(0.473)	(0.472)
Good subjective life quality (dummy variable)	0.899 (0.302)	0.912 (0.283)	0.888 (0.316)	0.900 (0.299)	0.896 (0.305)
B. Running Variable					
NCEE score	431.690	445.244	421.675	444.001	418.809
	(106.340)	(102.960)	(107.719)	(100.447)	(110.745)
Normalized NCEE score = (NCEE score - undergraduate score line)/10	-35.475 (97.080)	-21.857 (97.200)	-45.778 (95.770)	-21.919 (96.098)	-49.665 (96.096)

	Whole sample	**Urban residents**	**Rural residents**	**Science subject**	**Arts subject**
C. Control variables:					
Urban residence registration	0.424			0.420	0.429
	(0.494)			(0.494)	(0.495)
Age	30.749	31.374	30.272	30.053	31.469
	(11.288)	(11.229)	(11.307)	(10.818)	(11.713)
D. Predetermined variables:					
Province	36.693	35.095	37.796	36.481	36.831
	(14.230)	(15.292)	(13.255)	(14.552)	(13.892)
Father's occupation	6.362	4.951	7.407	6.415	6.307
	(2.733)	(2.875)	(2.080)	(2.722)	(2.745)
Mother's occupation	7.232	6.169	8.022	7.287	7.174
	(2.265)	(2.801)	(1.286)	(2.219)	(2.311)
Parents' education	2.269	2.825	1.859	2.288	2.250
	(1.559)	(1.866)	(1.127)	(1.598)	(1.518)
The only child	0.309	0.464	0.194	0.315	0.303
	(0.462)	(0.499)	(0.395)	(0.464)	(0.459)
Female	0.464	0.471	0.460	0.360	0.572
	(0.499)	(0.499)	(0.498)	(0.480)	(0.495)
Science subject	0.508	0.503	0.512		
	(0.500)	(0.500)	(0.500)		
Han nationality	0.939	0.938	0.939	0.945	0.932
	(0.240)	(0.241)	(0.239)	(0.228)	(0.252)
Observations	4002	1691	2295	2034	1968

4. Results and Discussion

4.1. First Stage Results: Testing If Passing the Score-line a valid IV Estimator

Table 2 mainly shows the first stage results, which is the effect of passing the undergraduate admission score line on the probability of getting an undergraduate degree (i.e., the probability of getting the treatment). In Table 2, overall, the effect of passing the score line on the probability of getting an undergraduate degree is estimated to be statistically significant and positive. For individuals passing the score line, though its effect on getting an undergraduate degree is in the whole sample, it is not necessarily that every individual passing the score line can earn an undergraduate degree. Hence, this technically makes this a case of Fuzzy RD Design. However, it is noticeable that the estimated coefficient of passing the score line in urban region is insignificant. This is probably because people living in cities tend to choose more competitive majors rather than better universities. Also, in the arts subject group, the estimated effect is also insignificant. Therefore, to comprehensively estimate the effect of better academic performance, both sharp RDD and fuzzy RDD methods are used. The sharp RDD uses the normalized score line as the cutoff point, while the fuzzy RDD uses passing the score line as an IV variable.

Table 2. Effect of passing undergraduate admission score line on the probability of getting an undergraduate degree

Variables	Whole undergraduate	Urban undergraduate	Rural undergraduate	Science undergraduate	Arts undergraduate
Passing the score line	0.108***	0.004	0.181***	0.128**	0.085
	(0.040)	(0.046)	(0.050)	(0.051)	(0.053)
Normalized score/10	0.039***	0.045***	0.034***	0.040***	0.036***
	(0.005)	(0.006)	(0.006)	(0.007)	(0.004)
Passing the score line * (Normalized score/10)	-0.017**	-0.015	-0.017*	-0.016*	-0.015
	(0.007)	(0.010)	(0.009)	(0.009)	(0.011)
(Normalized score/10)^2	0.001***	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Passing the score line * (Normalized score/10)^2	-0.001***	-0.002***	-0.001**	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Han nationality	-0.024	-0.016	-0.012	0.041	-0.068
	(0.027)	(0.043)	(0.040)	(0.044)	(0.044)
Science subject	0.011	0.036	-0.004		
	(0.014)	(0.025)	(0.020)		
Age	0.015**	0.007	0.026**	0.015	0.022**
	(0.007)	(0.012)	(0.010)	(0.013)	(0.010)
Age^2	-0.000*	-0.000	-0.000***	-0.000	-0.000**
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance exam controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	0.405	0.511	0.327	0.453	0.355
Observations	3729	1591	2125	1899	1817
R-squared	0.211	0.303	0.168	0.27	0.191

4.2. Reduced Form – Second Stage of Fuzzy RDD

The results of reduced form are displayed from Table 3 to Table 6. The estimations of the undergraduate variable in second stage regressions are also the main interests of this research, estimating whether an undergraduate degree impacts labor outcomes and life quality. Table 3 and Table 4 display the estimated effect on labor outcomes. As for the estimated effect on income (Table 3), no significant effect is discovered in most variables related to test scores. However, age is estimated to positively affect log income. Also, in Table 4, the effect of test score on probability of getting employed is not significant, while age's impact is statistically significant and positive. The estimated results in Table 3 and Table 4 indicate that working experience is more likely to be an important determinant of labor outcomes.

Table 5 and Table 6 display the estimated effect on life quality. In case of happiness (Table 5), all controlled variables are not estimated to have significant effect on happiness. Table 6 shows that getting an undergraduate degree seems not to improve the subjective life quality in most cases. Since most of the estimators in Table 5 and Table 6 are insignificant, which indicates that the determinants of life quality are different from those of labor outcomes. Instead, health, family, and marriage are more likely to affect life quality. Therefore, the results of second stage show that getting an undergraduate degree does not directly determine labor outcomes and life quality.

Table 3. Effect of getting an undergraduate degree on annual income

Variables	Whole	Urban	Rural	Science	Arts
	log(income)	log(income)	log(income)	log(income)	log(income)
Undergraduate degree	-0.187	-5.329	0.082	-0.137	-0.269
	(0.466)	(16.957)	(0.406)	(0.589)	(0.784)
Normalized score/10	0.015	0.292	-0.001	0.001	0.026
	(0.023)	(0.867)	(0.019)	(0.031)	(0.035)
Passing the score line * (Normalized score/10)	0.019	-0.15	0.024	0.048**	-0.009
	(0.018)	(0.493)	(0.019)	(0.021)	(0.033)
(Normalized score/10)^2	0.000	0.006	-0.000	-0.000	0.001
	(0.001)	(0.018)	(0.000)	(0.001)	(0.001)
Passing the score line * (Normalized score/10)^2	-0.001	-0.011	0.000	-0.001	-0.000
	(0.001)	(0.032)	(0.001)	(0.001)	(0.001)
Han nationality	0.122*	0.183	0.076	0.111	0.116
	(0.074)	(0.381)	(0.092)	(0.087)	(0.153)
Science subject	0.060*	0.476	0.016		
	(0.033)	(1.176)	(0.041)		
Age	0.081***	0.128	0.092**	0.120***	0.072**
	(0.027)	(0.300)	(0.037)	(0.046)	(0.034)
Age^2	-0.001***	-0.001	-0.001**	-0.002***	-0.001*
	(0.000)	(0.003)	(0.000)	(0.001)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	10.342	10.493	10.221	10.391	10.294
Observations	2373	1068	1295	1175	1188

Table 4. Effect of getting an undergraduate degree on employment

Variables	**Whole employed**	**Urban employed**	**Rural employed**	**Science employed**	**Arts employed**
Undergraduate degree	-0.041	0.254	-0.074	0.085	-0.205
	(0.212)	(8.811)	(0.170)	(0.259)	(0.383)
Normalized score/10	-0.004	-0.011	-0.005	-0.013	0.005
	(0.010)	(0.396)	(0.008)	(0.014)	(0.016)
Passing the score line * (Normalized score/10)	0.002	0.003	0.000	0.002	0.006
	(0.006)	(0.135)	(0.007)	(0.008)	(0.008)
(Normalized score/10)^2	-0.000	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.008)	(0.000)	(0.000)	(0.000)
Passing the score line * (Normalized score/10)^2	0.000	0.000	0.000	0.001	-0.001
	(0.000)	(0.018)	(0.000)	(0.001)	(0.001)
Han nationality	0.023	0.036	0.018	-0.003	0.025
	(0.024)	(0.143)	(0.031)	(0.037)	(0.044)
Science subject	-0.015	-0.045	0.001		
	(0.011)	(0.321)	(0.016)		
Age	0.081***	0.083	0.068***	0.083***	0.078***
	(0.009)	(0.064)	(0.013)	(0.014)	(0.015)
Age^2	-0.001***	-0.001	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	0.652	0.668	0.641	0.631	0.674
Observations	3729	1591	2125	1899	1817

Table 5 and Table 6 display the estimated effect on life quality. In the case of happiness (Table 5), all controlled variables are not estimated to have significant effect on happiness. Table 6 shows that getting an undergraduate degree seems to not improve subjective life quality in most cases. Since most of the estimators in Table 5 and Table 6 are insignificant, it indicates that the determinants of life quality are different from those of labor outcomes. Instead, health, family, and marriage are more likely to affect life quality. Therefore, the results of second stage show that getting an undergraduate degree does not directly determine labor outcomes and life quality.

Table 5. Effect of getting an undergraduate degree on level of happiness

Variables	Whole happy	Urban happy	Rural happy	Science happy	Arts happy
Undergraduate degree	0.308	23.114	0.162	0.615	-0.113
	(0.326)	(740.530)	(0.249)	(0.408)	(0.632)
Normalized score/10	-0.016	-1.033	-0.006	-0.038*	0.007
	(0.016)	(32.939)	(0.012)	(0.021)	(0.027)
Passing the score line * (Normalized score/10)	0.010	0.367	-0.006	0.023*	-0.006
	(0.009)	(11.224)	(0.010)	(0.012)	(0.013)
(Normalized score/10)^2	-0.000	-0.021	-0.000	-0.001*	0.000
	(0.000)	(0.653)	(0.000)	(0.001)	(0.001)
Passing the score line * (Normalized score/10)^2	0.001	0.047	0.001*	0.002	0.000
	(0.001)	(1.515)	(0.001)	(0.001)	(0.001)
Han nationality	0.003	0.271	-0.004	-0.021	-0.009
	(0.035)	(8.950)	(0.048)	(0.059)	(0.059)
Science subject	-0.008	-0.913	0.001		
	(0.016)	(28.962)	(0.021)		
Age	0.006	-0.123	0.014	0.007	0.009
	(0.012)	(3.947)	(0.017)	(0.018)	(0.019)
Age^2	-0.000	0.001	-0.000	-0.000	-0.000
	(0.000)	(0.042)	(0.000)	(0.000)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	0.664	0.676	0.657	0.663	0.665
Observations	3680	1566	2101	1872	1795

Table 6. Effect of getting an undergraduate degree on subjective life quality

Variables	Whole life quality	Urban life quality	Rural life quality	Science life quality	Arts life quality
Undergraduate degree	-0.192	-18.325	-0.089	-0.097	-0.361
	(0.202)	(589.344)	(0.159)	(0.214)	(0.455)
Normalized score/10	0.012	0.817	0.008	0.009	0.017
	(0.009)	(26.214)	(0.007)	(0.011)	(0.019)
Passing the score line * (Normalized score/10)	-0.007	-0.279	-0.007	-0.007	-0.007
	(0.005)	(8.932)	(0.006)	(0.006)	(0.008)
(Normalized score/10)^2	0.000	0.016	0.000	0.000	0.000
	(0.000)	(0.519)	(0.000)	(0.000)	(0.000)
Passing the score line * (Normalized score/10)^2	-0.000	-0.037	-0.000	-0.000	-0.001
	(0.000)	(1.206)	(0.000)	(0.000)	(0.001)
Han nationality	0.042*	-0.182	0.050	0.076*	0.005
	(0.025)	(7.119)	(0.033)	(0.039)	(0.043)
Science subject	0.003	0.719	0.002		
	(0.011)	(23.048)	(0.014)		
Age	0.000	0.100	0.003	-0.010	0.015
	(0.007)	(3.141)	(0.013)	(0.009)	(0.013)
Age^2	-0.000	-0.001	-0.000	0.000	-0.000
	(0.000)	(0.033)	(0.000)	(0.000)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	0.899	0.912	0.888	0.900	0.896
Observations	3679	1566	2100	1872	1794

4.3. Sharp RDD: Passing Score Line's Effect on Labor Outcomes and Life Quality

Since the Fuzzy RDD model, which is first supposed to be used, does not indicate that an undergraduate degree will affect labor outcomes and life quality, this research moves on to sharp RDD to analyze whether passing the undergraduate admission score line impacts those outcomes. The results of sharp RDD can be seen from Table 7 to Table 10, with each Table presenting results for one outcome. The labor outcomes are presented in Table 7 and Table 8. The outcomes related to happiness, which are selected as level of happiness and subjective life quality, are presented in Table 9 and Table 10.

According to the estimation results shown in Table 7 and Table 8, the effect of age is statistically significant and positive for labor outcomes, instead of test scores. As for the income case shown in Table 7, in most groups, for individuals who pass the score line, the estimated effect of increasing test scores does not have a significant effect on the log(income). However, in the science subsample, for individuals who pass the score line, the estimated effect of increasing test scores by 10 points is statistically significant and positive. It is noticeable that in the whole sample and all defined subsamples, age's effect on income is estimated to be statistically significant and positive. Even with the marginal decreasing effect, the effect is still positive for individuals in the working age. Among the employment rate estimators for different groups shown in Table 8, for individuals in all groups who pass the score line, the estimated effect of increasing test scores does not have a significant effect on the possibility of having a job.

Moreover, most estimators in Table 8 related to test scores are not statistically significant. Like the estimated results for log(income), in the whole sample and all subsamples, age's effect on employment is estimated to be statistically significant and positive. This result is understandable since individuals with more working experience tend to be more competitive in the labor market, leading to higher income or more employment opportunities. Also, during the survey, some young individuals were on their way to job-seeking or studying.

Table 7. Effect of passing undergraduate admission score line on yearly income

Variables	Whole	Urban	Rural	Science	Arts
	log(income)	log(income)	log(income)	log(income)	log(income)
Passing the score line	-0.024	-0.104	0.017	-0.02	-0.027
	(0.051)	(0.106)	(0.058)	(0.064)	(0.081)
Normalized score/10	0.008	0.028**	0.002	-0.005	0.016
	(0.007)	(0.013)	(0.008)	(0.008)	(0.01)
Passing the score line * (Normalized score/10)	0.024*	0.005	0.022	0.051***	-0.000
	(0.014)	(0.02)	(0.015)	(0.014)	(0.024)
(Normalized score/10)^2	-0.000	0.001	-0.000	-0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Passing the score line * (Normalized score/10)^2	-0.000	-0.001*	-0.000	-0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Han nationality	0.131	0.225*	0.071	0.101	0.152
	(0.079)	(0.114)	(0.093)	(0.067)	(0.122)
Science subject	0.052*	0.107***	0.018		
	(0.030)	(0.034)	(0.040)		
Age	0.078**	0.041	0.094**	0.116**	0.068**
	(0.031)	(0.042)	(0.046)	(0.053)	(0.033)
Age^2	-0.001**	-0.000	-0.001**	-0.002**	-0.001*
	(0.000)	(0.001)	(0.001)	(0.001)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	10.342	10.493	10.221	10.391	10.294
R-squared	0.211	0.303	0.168	0.27	0.191
Observations	2373	1068	1295	1175	1188

Table 8. Effect of passing undergraduate admission score line on employment

Variables	Whole employed	Urban employed	Rural employed	Science employed	Arts employed
Passing the score line	-0.004	0.001	-0.013	0.011	-0.017
	(0.015)	(0.027)	(0.024)	(0.025)	(0.029)
Normalized score/10	-0.006**	0.001	-0.008**	-0.010*	-0.003
	(0.003)	(0.005)	(0.004)	(0.005)	(0.003)
Passing the score line * (Normalized score/10)	0.003	-0.001	0.001	0.001	0.009
	(0.004)	(0.006)	(0.005)	(0.007)	(0.006)
(Normalized score/10)^2	-0.000	0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Passing the score line * (Normalized score/10)^2	0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Han nationality	0.024	0.032	0.019	0.001	0.039
	(0.027)	(0.046)	(0.026)	(0.043)	(0037)
Science subject	-0.016	-0.036*	0.001		
	(0.012)	(0.020)	(0.013)		
Age	0.081***	0.085***	0.066***	0.084***	0.074***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Age^2	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	0.652	0.668	0.641	0.631	0.674
R-squared	0.513	0.578	0.488	0.544	0.495
Observations	3729	1591	2125	1899	1817

This study uses level of happiness and subjective life quality as measurements of life quality. Table 9 and Table 10 show the estimation results. According to the estimation results, the values of estimators show that most of the endogenous valuables controlled in this model do not have any significant impact on the level of happiness, which is identical among all groups. The results indicate that happiness can be affected by factors other than academic performance. Instead of test scores, happiness can be affected by such factors as health, family atmosphere, and marriage. Also, since both measurements are self-reported, they should be influenced by many factors in individuals' daily lives.

Table 9. Effect of passing undergraduate admission score line on happiness

Variables	**Whole happiness**	**Urban happiness**	**Rural happiness**	**Science happiness**	**Arts happiness**
Passing the score line	0.032	0.035	0.029	0.080	-0.009
	(0.026)	(0.059)	(0.024)	(0.052)	(0.056)
Normalized score/10	-0.004	-0.007	-0.001	-0.014**	0.002
	(0.004)	(0.008)	(0.004)	(0.006)	(0.005)
Passing the score line * (Normalized score/10)	0.005	0.017*	-0.009	0.014	-0.004
	(0.006)	(0.010)	(0.007)	(0.009)	(0.012)
(Normalized score/10)^2	-0.000*	-0.000	-0.000	-0.000**	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Passing the score line * (Normalized score/10)^2	0.000*	-0.000	0.001**	0.001*	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Han nationality	-0.003	-0.008	-0.005	0.006	-0.003
	(0.028)	(0.051)	(0.047)	(0.051)	(0.046)
Science subject	-0.005	-0.008	-0.000		
	(0.017)	(0.029)	(0.022)		
Age	0.011	0.001	0.018	0.016	0.007
	(0.011)	(0.018)	(0.013)	(0.013)	(0.015)
Age^2	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	0.664	0.676	0.657	0.663	0.665
R-squared	0.033	0.051	0.041	0.044	0.051
Observations	3680	1566	2101	1872	1795

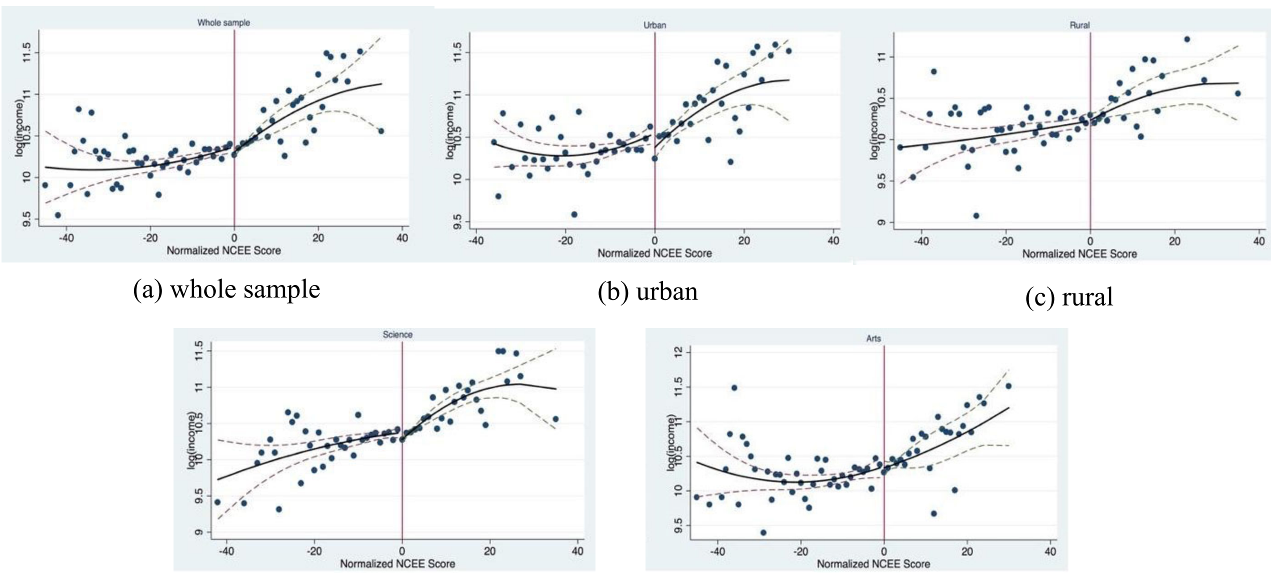
Table 10. Effect of passing undergraduate admission score line on happiness

Variables	**Whole life quality**	**Urban life quality**	**Rural life quality**	**Science life quality**	**Arts life quality**
Passing the score line	-0.020	-0.028	-0.016	-0.013	-0.028
	(0.021)	(0.038)	(0.026)	(0.028)	(0.028)
Normalized score/10	0.005*	0.004	0.006*	0.005*	0.004
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Passing the score line * (Normalized score/10)	-0.004	-0.001	-0.006	-0.006	-0.002
	(0.004)	(0.007)	(0.005)	(0.006)	(0.005)
(Normalized score/10)^2	0.000	0.000	0.000*	0.000**	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Passing the score line * (Normalized score/10)^2	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Han nationality	0.046**	0.040	0.051*	0.071**	0.026
	(0.022)	(0.036)	(0.028)	(0.033)	(0.031)

Variables	Whole life quality	Urban life quality	Rural life quality	Science life quality	Arts life quality
Science subject	0.001	0.001	0.002		
	(0.010)	(0.018)	(0.011)		
Age	-0.002	0.002	0.001	-0.012	0.008
	(0.005)	(0.007)	(0.012)	(0.008)	(0.010)
Age^2	0.000	0.000	-0.000	0.000*	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Province controls	Yes	Yes	Yes	Yes	Yes
Year of taking college entrance examination controls	Yes	Yes	Yes	Yes	Yes
Standard errors	Cluster	Cluster	Cluster	Cluster	Cluster
Mean	0.899	0.912	0.888	0.900	0.896
R-squared	0.048	0.064	0.054	0.063	0.059
Observations	3679	1566	2100	1872	1794

4.4. Showing Regression Discontinuity Using Graphs

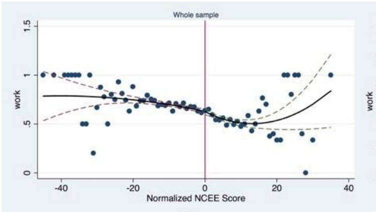
The Sharp RDD's graphs depicting the score's effect on labor outcomes and life quality are shown from Fig.1 to Fig.4. The labor outcomes are depicted in Fig.1 and Fig.2. Fig.1 clearly shows that individuals with higher scores will earn more, no matter they pass the score line or not. In Fig.1, in the urban subsample (Fig.1 (b)) and science subject subsample (Fig.1 (d)), there is a kink on just the right side of cutoff, with a marginal decreasing effect. No noticeable kink of slope is detected in other graphs. This indicates that for urban residents' and science students' who pass the score line, the marginal effect of score becomes higher than those below the score line. This shows that a higher score can increase their earnings with more potential. Nevertheless, the effect becomes smaller when the score becomes higher. In Fig.2, there is no noticeable change of slope when passing the cutoff. Instead, the slope even does not change, indicating that passing the score line itself does not have large impact on the chance of getting employed. For individuals just around the cutoff, the slope is even negative. However, this cannot indicate a negative effect on employment since the magnitude of slope is small in general.



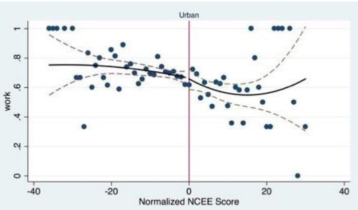
(a) whole sample (b) urban (c) rural

(d) science (e) arts

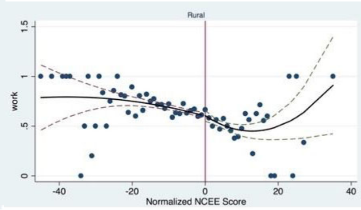
Fig. 1 Effect of passing undergraduate admission score line on yearly income.



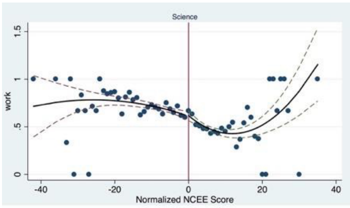
(a) whole sample



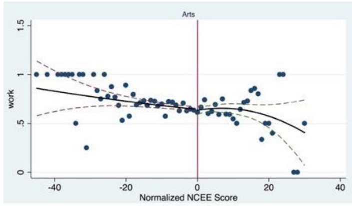
(b) urban



(c) rural



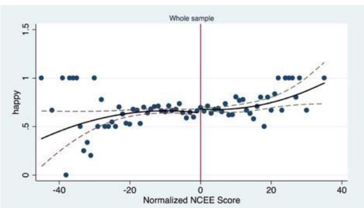
(d) science



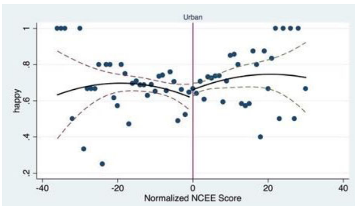
(e) arts

Fig. 2 Effect of passing undergraduate admission score line on employment.

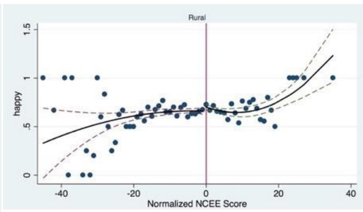
The life-quality outcomes are depicted in Fig.3 and Fig.4. Fig.3 generally shows that test score positively affects happiness level, but the effect is small considering the slope. However, in Fig.3, for most groups, there is no jump or detectable change of slope around the cutoff. Though in the urban subsample (Fig.3(b)), there seems to be a detectable slope change around the cutoff, the change is subtle since the magnitude of the change is small. In the urban subsample, the effect could be unclear, indicating that their happiness level is affected by various factors. In Fig.4, which depicts the effect of test scores on subjective life quality, the slope changes around the cutoff in all groups are minimal, and the slopes remain small among all the groups. Moreover, in Fig.3 and Fig.4, no jump is detected around the cutoff. Thus, Fig.3 and Fig.4 imply that passing the score line or higher test scores do not play an essential role in life quality.



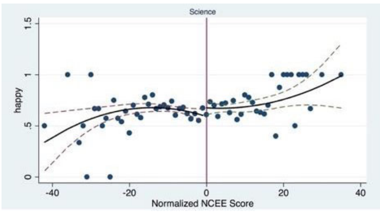
(a) whole sample



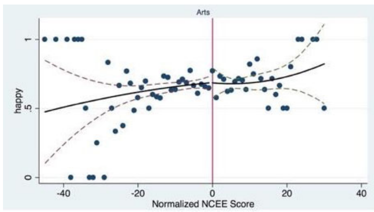
(b) urban



(c) rural

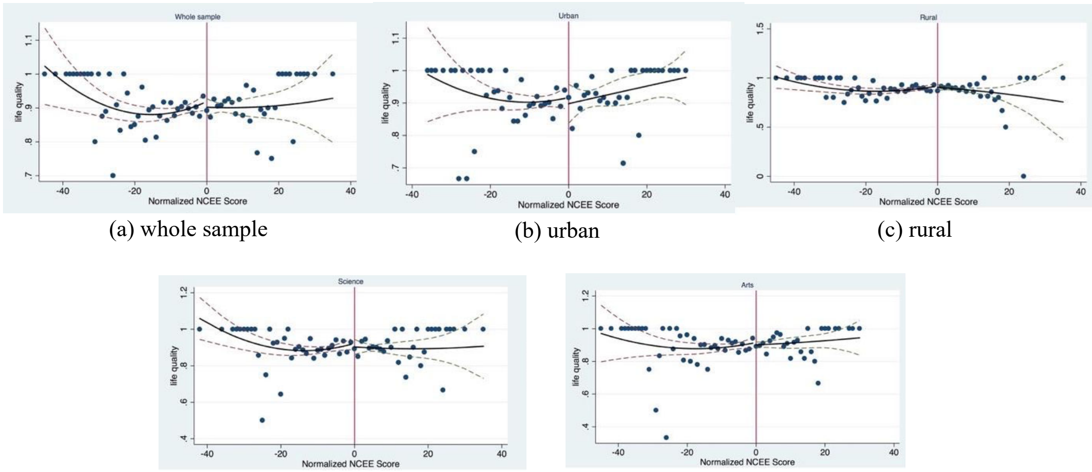


(d) science



(e) arts

Fig.3 Effect of passing undergraduate admission score line on level of happiness.



(a) whole sample (b) urban (c) rural

(d) science (e) arts

Fig.4 Effect of passing undergraduate admission score line on subjective life quality

Among all graphs, even if there is a detectable change around the cutoff, most of the change around the cutoff is a small-magnitude kink, not a jump. Most graphs show no detectable change of slope. The visualized results indicate that though higher test scores can make people earn more or become happier, the effects caused by passing the score line are subtle.

4.5. Discussion and Explanation of Results

Overall, the results derived from this study are quite counterintuitive. Though the results show that a better score in the college entrance examination can make people richer and happier, the effects are weakly related to passing the score line or getting an undergraduate degree. Therefore, the results are only partially consistent with the previous works, most of which generally showed a high and usually significant educational return if people get a better degree or pass a certain score line.

However, it is noticeable that higher scores may have a positive effect in some groups. A typical case is the income of science subsample. In both fuzzy and sharp RDD methods, for individuals in the science subsample who pass the score line, the effect of a ten-point increase on income is estimated to be significant and positive. This can show that the nature of subjects influences the effect of passing the score line. Works related to science subjects usually require employees with better learning abilities and creativity, so students with better academic performance tend to be more competitive.

The findings also uncover other factors that influence labor outcomes and life quality. For the labor outcomes, three factors can explain the findings:

(1) Older people tend to have more work experience, which means they are likely to get employed and earn more.

(2) Some policies can influence individuals' likelihood of pursuing an undergraduate degree. One typical example is the college admission extension policy released in 1999, which brought more chances for individuals to enter universities. The policy means that a university degree has become common, and a university degree has been less competitive since then.

(3) In the Chinese labor market, where the survey is done, advantages like job certification, social ability, and so on are more important in many occupations.

For life quality, test scores and working experience do not have significant effect on life quality because the determinants of life quality are different from those of labor outcomes.

5. Conclusion

The main finding of this research is that a better score in the college entrance examination may have positive effect on income and life quality. However, the effects of passing the undergraduate admission score line and getting an undergraduate degree are estimated to be subtle. Therefore, this study has several indications. First, individuals with better academic performance have the potential to become richer, no matter whether they pass the score line or get an undergraduate degree or not – this finding from sharp RDD method is similar to previous studies using OLS method. Also, in the Chinese labor market, abilities other than academic performance make people more competitive. Such factors as social ability, networking, and certifications usually make it easier to get a job and give people more chances to earn more. Those factors are usually tightly connected with working experience. This study also implies that life quality is likely to be determined by factors other than academic performance and working experience.

Although the results derived from RDD do not show the importance of passing the score line or getting an undergraduate degree, this study somewhat serves to re-think the power of undergraduate education in the Chinese labor market. The findings also have some educational policy implications in China. Notably, the courses should be more connected with the jobs students may do in the future. Besides, universities can include courses such as social networking, career planning, and methods to cope with mental stress to prepare students for more competitive occupations and better life quality.

References

[1] Nigel C. O'Leary, Peter J. Sloane. 2005. The Return to a University Education in Great Britain. National Institute Economic Review193, No.1:75-89.

[2] LUO Chuliang. 2007. Return to Education and Its Distribution in Urban China. Economic Research Jounal 6:119-129.

[3] Elliott Fan, Xin Meng, Zhichao Wei, and Guochang Zhao. 2010. Rates of Return to University Education: The Regression Discontinuity Design. IZA Discussion Paper, No. 4749.

[4] LIU Zeyun. 2015. Long-term Change in Returns to Higher Education in China: 1988-2007. Peking University Education Review 13. No.14.:65-81.

[5] Arjun S. Bedi, Noel Gaston. 1998. Using variation in schooling availability to estimate educational returns for Honduras. Economics of Education Review 18: 107-116.

[6] Dong Yingying. 2018. Jump or Kink? Regression Probability Jump and Kink Design for Treatment Effect Evaluation. Working paper, University of California, Irvine.

[7] Prashant Loyalka, Yingquan Song, Jianguo Wei. 2012. The effects of attending selective college tiers in China. Social Science Research 41: 287-305.

[8] Jere R. Behrman, Yanyan Xiong, Junsen Zhang Cross-sectional schooling-health associations misrepresented causal schooling effects on adult health and health-related behaviors: Evidence from the Chinese Adults Twins Survey. Soc Sci Med. 2015 Feb; 127: 190–197.

[9] Juncal Cuñado, Fernando Pérez de Gracia.Does Education Affect Happiness? Evidence for Spain. 2012. Social Indicators Research volume 108, pages185–196 (2012).

[10] Bin Huang, Massimiliano Tani, Yi Wei, Yu Zhu. 2022. Returns to education in China: Evidence from the great higher education expansion. China Economic Review: 1-22.