Journal of Economics
and Administrative Sciences (JEAS)

Available online at http://jeasiq.uobaghdad.edu.iq

Estimation of Causal Effect of treatment via Fuzzy Regression Discontinuity Designs

Sajad Sammer Abd Al-Razzaq
Department of Statistics,
College of Administration & Economics,
University of Baghdad, Iraq
Sajad.samir1201a@coadec.uobaghdad.edu.iq

Mohammad Jasim Mohammad
Department of Statistics,
College of Administration & Economics,
University of Baghdad, Iraq
m.jasim@coadec.uobaghdad.edu.iq

Received:13/11/2022 Accepted:4/12/2022 Published: December / 2022

Abstract

In some cases, researchers need to know the causal effect of the treatment in order to know the extent of the effect of the treatment on the sample in order to continue to give the treatment or stop the treatment because it is of no use. The local weighted least squares method was used to estimate the parameters of the fuzzy regression discontinuous model, and the local polynomial method was used to estimate the bandwidth. Data were generated with sample sizes (75,100,125,150) in repetition 1000 .An experiment was conducted at the Innovation Institute for remedial lessons in 2021 for 72 students participating in the institute and data collection. Those who used the treatment had an increase in their score after treatment by 25.95%.

Paper type This Research is extracted from a master's thesis on statistics entitled "Estimation of Fuzzy Regression Discontinuity model with application"

Keywords: **Fuzzy Regression Discontinuity, Causal effect of the treatment, Robust Local Polynomial Regression Estimators, Coverage Error Optimality.**

1. Introduction

The regression discontinuous design was first introduced by (Thistlethwaite and Campbell, 1960). It was presented to find out the value of the causal effect of treatment. Many economic and other social science experiments rely on program causal effects. The regression discontinuous design determines the causal effect of treatment by comparing the results of treated and untreated individuals near the cut-off point. It is assumed that those individuals directly below the cut-off point have the same unobserved distribution as those individuals directly above the cut-off point. The causal effect of the treatment will be estimated by the data obtained from the trial.

Fuzzy groups have great flexibility in describing many real cases, because they identify the people who received treatment and the people who did not receive treatment (Rani,2014), where its boundaries are not lucid between people who take treatment and people who do not take treatment. (Ali,2013)

Used (Angrist and Rokkanen,2015) fuzzy regression discontinued to study the effects of Boston exam schools for applicants who apply to an area below the cut-off point (acceptance point). These estimates indicate that the causal effects of school enrollment on the exam for applicants in ninth grade who have values that are very far from the cut-off point (acceptance point) differ slightly from those for applicants with values that put them on the fringe of the cut-off point (acceptance point).

There are many problems facing education in Iraq, and they have several reasons, including the student's unwillingness to study, poor education in schools, or an increase in the number of students in schools. An experiment was made for the students of the Innovation Institute for remedial lessons before registering at the institute and after giving them a lecture to find out the causal effect of the lecture in order to find out the extent of the student's desire to study.

2. Materials and Methods

2.1 Fuzzy regression discontinuity

In the fuzzy regression discontinuity design, the treatment effect is determined by a discontinuity in the conditional probability of treatment. The confusion may result due to the incomplete implementation that treated some of the items that are not eligible for treatment or neglected the treatment of some eligible items or treated some of the unqualified items and neglected some of the eligible items. (Angrist and Imbens, 1994). In some cases, members receive treatment even though they do not need treatment because their value is higher than the cutoff set by the researcher, and it can be written as follows: (Imbens, G. W., and T. Lemieux, 2007)

$$\lim_{x\downarrow c} \Pr(T_i = 1|X_i = x) \neq \lim_{x\uparrow c} \Pr(T_i = 1|X_i = x) \qquad \dots(1)$$

T_i: Treatment

X_i: Sample values

Since it is impossible to measure T for the individual himself when he takes treatment and when he does not take treatment, and thus it is impossible to measure the effect of treatment on the individual outcome and therefore we will measure the average local causal effect of the treatment for the community. (Yang He and Otávio Bartalotti,2020). It can be written in the following form:

$$\mathcal{T} = \frac{\lim\limits_{x\downarrow c} \mathbb{E}(\mathbf{W}|\mathbf{X}=\mathbf{c}) - \lim\limits_{x\uparrow c} \mathbb{E}(\mathbf{W}|\mathbf{X}=\mathbf{c})}{\lim\limits_{x\downarrow c} \mathbb{E}(\mathbf{T}|\mathbf{X}=\mathbf{c}) - \lim\limits_{x\uparrow c} \mathbb{E}(\mathbf{T}|\mathbf{X}=\mathbf{c})} \quad \dots (2)$$

$\mathcal{T}$: The causal effect of the treatment.
c: cut point.
W: dependent variable.

The numerator measures the jump in the result at cutting, and the denominator measures the jump in processing levels at the cut and provides us with an estimate of the average effect. Where $T_i(c)$ a possible treatment state at cut-off point c, and $T_i(c)$ is one if the unit is receiving treatment if the cut-off point c. There are two models, parametric and non-parametric for fuzzy regression discontinuous.

We will use the below linear regression discontinuous model can be approved linear fuzzy regression discontinuous model Depending on the probability of receiving the treatment :

$$W_i = \alpha_0 + \mathcal{T} T_i + \alpha_1 d_i + \varepsilon_i \quad \cdots (3)$$

Where
W_i:The response variable (post-trial).
α_0: The constant limit parameter.
$\mathcal{T}$: The parameter of the causal effect of the treatment.
T_i: The classification variable (ordinal variable) and it is the explanatory variable in this model.
α_1: The parameter of the explanatory variable X_i .
$d_i = (X_i - c)$.
ε_i: The error term of the model, which is represented randomly and distributed normally with zero mean and variance σ^2.

2.2 Methods of Estimation

2.2.1 Robust Local Polynomial Regression Estimators

It is a nonparametric estimation method for estimating the bandwidth parameter of the fuzzy regression discontinuous gradient ,and it is one of the preferred options for the nonparametric estimators in the fuzzy regression discontinuous literature due to its good boundary properties (Fan and Gijbels, 1996).

An important characteristic of this method is the choice of polynomial order and bandwidth, both of which have a significant impact on the estimators, because if not chosen correctly will lead to incorrect estimates of the model parameters. (Yang He, 2017)

We will estimate the bandwidth using the coverage error optimality method. (Calonico,et all,2017)

$$\hat{h}_{cer} = n^{-\frac{p}{(3+p)(3+2p)}} \hat{h}_{MSE} \quad \cdots (4)$$

Which is based on the following fuzzy regression discontinuous estimations:
1- The estimation of the local causal effect of the treatment is as follows:

$$h_{cermserd}(\widehat{\varphi}\mathrm{h}) = \widehat{\mathcal{T}} = \frac{\widehat{\mathcal{T}}_W}{\mathcal{T}_T} \quad \cdots (5)$$

2- Amount of causal effect on the left side:

$$hcer_{left}(\widehat{\varphi}\mathrm{h}) = \widehat{\mathcal{T}}_{Wl} = \widehat{\alpha}_{0l} \quad \cdots (6)$$

3- Amount of causal effect on the right side:

$$h_{cerright}(\widehat{\varphi}\mathrm{h}) = \widehat{\mathcal{T}}_{Wr} = \widehat{\alpha}_{0r} \quad \cdots (7)$$

4- Estimated the sum of the local causal effect of the regression:

$$h_{cersum}(\widehat{\varphi}\mathrm{h}) = \widehat{\mathcal{T}}_W = \widehat{\mu}_{Wr} + \widehat{\mu}_{Wl} \quad \cdots (8)$$

After using the above four estimators and estimating the bandwidth, we will perform some calculations on them and find the following bandwidth estimators:
1- We will take the lowest estimate of the bandwidth from the estimator of the local causal effect of the treatment and the estimator of the sum of the local causal effect of the regression as follows:

$$\widehat{h}_{cercomb1=} min(\widehat{h}_{cermserd}, \widehat{h}_{cersum}) \quad \cdots (9)$$

2 – We will take the median of the bandwidth estimators from the local causal effect estimator of the treatment, the left side causal effect estimator, and the right side causal effect estimator, the sum of the local causal effect estimator for the regression, as follows: (Calonico,et all,2017)

$$\widehat{h}_{cercomb2=} median(\widehat{h}_{cermserd}, \widehat{h}_{cersum}, \widehat{h}_{certwo}) \quad \cdots (10)$$

$$\widehat{h}_{certwo} = (\widehat{h}_{cerleft}, \widehat{h}_{cerright}) \quad \cdots (11)$$

2.3.2 Local weighted least squares method

We will use the local weighted least squares method to estimate parameters $\widehat{\alpha}_{0l}$, $\widehat{\alpha}_{0r}$, $\widehat{\beta}_{0l}$, $\widehat{\beta}_{0r}$, $\widehat{\alpha}_{1l}$, $\widehat{\alpha}_{1r}$, $\widehat{\beta}_{1l}$, $\widehat{\beta}_{1r}$. (Donna et al,2012)

$$(\widehat{\alpha}_{0l}, \widehat{\alpha}_{1l}) = \min_{\alpha_1} \sum_{i=1}^{n} 1(X_i < c)\left(W_i - \alpha_{0l} - \alpha_{1l}(X_i - c)\right)^2 K_h\left(\frac{X_i - c}{h}\right) \quad \cdots (12)$$

$$(\widehat{\alpha}_{0r}, \widehat{\alpha}_{1r}) = \min_{\alpha_1} \sum_{i=1}^{n} 1(X_i \geq c)\left(W_i - \alpha_{0r} - \alpha_{1r}(X_i - c)\right)^2 K_h\left(\frac{X_i - c}{h}\right) \quad \cdots (13)$$

$$(\widehat{\beta}_{0l}, \widehat{\beta}_{1l}) = \min_{\beta_1} \sum_{i=1}^{n} 1(X_i < c)\left(T_i - \beta_{0l} - \beta_{1l}(X_i - c)\right)^2 K_h\left(\frac{X_i - c}{h}\right) \quad \cdots (14)$$

$$(\widehat{\beta}_{0r}, \widehat{\beta}_{1r}) = \min_{\beta_1} \sum_{i=1}^{n} 1(X_i \geq c)\left(T_i - \beta_{0r} - \beta_{1r}(X_i - c)\right)^2 K_h\left(\frac{X_i - c}{h}\right) \cdots \quad (15)$$

$\widehat{\alpha}_{0l}$: The estimator of the constant term parameter of the left side of the function W.
$\widehat{\alpha}_{0r}$: The estimator of the constant term parameter of the right side of the function W.
$\widehat{\alpha}_{1l}$: The estimator of the parameter variable X_i to the left side of the function W.
$\widehat{\alpha}_{1r}$: The estimator of the parameter variable X_i to the right side of the function W.
$\widehat{\beta}_{0l}$: The estimator of the constant term parameter of the left side of the function T.

$\widehat{\beta}_{0r}$: The estimator of the constant term parameter of the right side of the function T.
$\widehat{\beta}_{1l}$: The estimator of the parameter variable X_i to the left side of the function T.
$\widehat{\beta}_{1r}$: The estimator of the parameter variable X_i to the right side of the function T.
$K_h(.)$: It is a kernel function and we assume it as follows:

$$K(u) = 1(u < 0)K(-u) + 1(u \geq 0)K(u) \quad \cdots \quad (16)$$

Using the parameters estimated in the local weighted least squares method, we can estimate the average causal effect as follows:

$$\hat{\tau} = \frac{\widehat{\mathcal{T}}_W}{\widehat{\mathcal{T}}_T} \quad \cdots \quad (17)$$

$$\widehat{\mathcal{T}}_W = \hat{\mu}_{Wr} - \hat{\mu}_{Wl} \quad \cdots (18)$$

$$\widehat{\mathcal{T}}_T = \hat{\mu}_{Tr} - \hat{\mu}_{Tl} \quad \cdots (19)$$

$$\hat{\mu}_{Wr} = \widehat{\alpha}_{0r} \quad \cdots (20)$$

$$\hat{\mu}_{Wl} = \widehat{\alpha}_{0l} \quad \cdots (21)$$

$$\hat{\mu}_{Tr} = \widehat{\beta}_{0r} \quad \cdots (22)$$

$$\hat{\mu}_{Tl} = \widehat{\beta}_{0l} \quad \cdots (23)$$

We will use Epanechnikov kernel function:

$$k(u) = \begin{cases} \frac{3}{4}(1 - u^2) & |u| \leq 1 \\ 0 & o.w \end{cases} \quad \cdots (24)$$

3. Discussion of Results

3.1 Simulation

The Monte Carlo experiment will be used to generate the data and as follows:

$$W_i = T + \varepsilon_{W_i} \quad , i = 1, \ldots, n \quad \cdots \quad (25)$$

The treatment is defined as follows:

$$T_i = 1\{\varepsilon_{x_i} \leq 0\} \times 1\{x_i < 0\} + 1\{\varepsilon_{x_i} \leq 0\} \times 1\{x_i \geq 0\} \cdots (26)$$

The errors ε_{W_i} and ε_{x_i} are distributed in the common bivariate normal distribution as follows:

$$\begin{pmatrix} \varepsilon_{W_i} \\ \varepsilon_{x_i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho_1 \\ \rho_2 & 1 \end{pmatrix}\right) \quad \cdots (27)$$

The independent variable has a normal distribution as follows:

$$x_i \sim N(0, 1) \quad \cdots (28)$$

And that ρ_1 and ρ_2 are hypothetical values, and if we assume that the covariance and covariance matrix of errors are the same, we will assume the values as follows:

Table 1. Values (ρ_1,ρ_2)

ρ_1	ρ_2
0.5	0.5
0.75	0.99

We will repeat the generation 1000 times and the results will be calculated. (Rong Ma et al, 2016)
The program (R Projct) was used in writing the program for data analysis.

Table 2. Mean Saquer Error of Model for Local weighted least squares method

N	75		100		125		150	
	$\rho_1 = \rho_2 = 0.5$	$\rho_1 = 0.75$ $\rho_2 = 0.99$	$\rho_1 = \rho_2 = 0.5$	$\rho_1 = 0.75$ $\rho_2 = 0.99$	$\rho_1 = \rho_2 = 0.5$	$\rho_1 = 0.75$ $\rho_2 = 0.99$	$\rho_1 = \rho_2 = 0.5$	$\rho_1 = 0.75$ $\rho_2 = 0.99$
cerrd	0.7307618	0.5588209	0.7706251	0.5849685	0.7854602	0.5969877	0.7905289	0.6017573
certwo	0.7499605	0.5737859	0.7788047	0.5918734	0.7893558	0.6024506	0.7924243	0.6034771
cersum	0.7347276	0.5601735	0.7750029	0.5869163	0.7876774	0.5972313	0.7908142	0.6017713
cercomb1	0.7227721	0.5519606	0.7678667	0.582387	0.7812434	0.5940485	0.7867712	0.5990934
cercomb2	0.72877	0.5591855	0.7707949	0.5854424	0.783779	0.596358	0.7893338	0.600681

Cerrd: it is choosing the estimator of the local causal effect of the treatment as in Equation (5).
Certwo: it is choosing the causal effect estimator for the left side and the causal effect estimator for the right side as in the equation (6) and Equation (7).
Cersum: it is choosing the estimator of the total local causal effect of the regression as in Equation (8).
cercomb1: it is the choice of the least estimator of the estimators cersum and cerrd as in equation (9).
cercomb2: it is the choice of the median for the estimators cersum, cerrd and certwo as in Equation (10).

The mean square error increases with the increase of the sample size, because when the sample size increases, the bandwidth decreases, and the cercomb1 method has the lowest mean square error of the rest of the methods.

3.2 Application

Data were collected from the Innovation Institute for Special Teaching, where the scores of 72 students were collected before and after treatment, and the method was used cercomb1 in finding results.

Table 3. Estimated values of model parameters when cercomb1

Cof.	$\hat{\alpha}_0$	$\hat{\alpha}_1$	$\hat{\mathcal{T}}$
Value	52.00095	0.12948960	25.95022

$$\widehat{W}_i = 52.00095 + 0.12948960 X_i + 25.95022 T_i$$

4. Conclusion

We notice that the mean squares of error increases with the increase in the sample size, due to the increase in the number of values that fall outside the bandwidth; we note that the minimum mean squared error for the bandwidth for all samples is cercomb1. And The causal effect of the treatment is 25.95. This means that individuals who received treatment increase their score after treatment by 25% . The relationship between the sample size and the mean squared error is an inverse relationship, because when the sample size increases, the bandwidth decreases.

References

1. Ali Mohammed and Mohammed Sadiq.(2013).”Design a multiple Fuzzy linear regression model for world oil prices”. University Of Baghdad

2. Angrist, Imbens.(1994).” Identification and Estimation of Local Average Treatment Effects”. Econometrica. Vol.62, No.2, PP. 467-475.

3. Angrist and Rokkanen.(2015).” Wanna Get Away? Regression Discontinuity Estimation of Exam School Effects Away From the Cutoff”. Journal of the American Statistical Association.Vol.110,No.512,PP 1331-1344.

4. Calonico, S., M. D. Cattaneo, and M. H. Farrell. (2017).” rdrobust: Software for regression-discontinuity designs”. The Stata Journal. Vol.17, No.2, PP. 372-404.

5. Donna Feir and Thomas Lemieux and Vadim Marmer. (2012). “Weak Identification in Fuzzy Regression Discontinuity Designs”.SSRN Electronic Journal. Vol.34, No 2, PP. 185-196.

6. Fan, J. and Gijbels, I. (1996).” Local Polynomial Modelling and Its Applications” .Chapman and Hall, London. Introductory Econometrics(4th Edition) A Modern Approach, by Jeffrey M. Wooldridge Hardcover, 888 Pages, Published 2008 by South-Western.

7. Imbens, G. W., and T. Lemieux.(2007). “Regression Discontinuity Designs: A Guide to Practice”. NATIONAL BUREAU OF ECONOMIC RESEARCH. Vol.142, No.2, PP. 615-635.

8. Rani Fouad and Mohammed Sadiq.(2014).”building fuzzy linear programming model with practical application in general company of oif oroducing”.University Of Baghdad

9. Rong Ma et al.(2016).” Robust Inference in Fuzzy Regression Discontinuity with Multiple Forcing Variables”. Central University of Finance and Economics.

10. Thistlethwaite, D. L., & Campbell, D. T. (1960)." Regression-discontinuity analysis: An alternative to the ex post facto experiment". *Journal of Educational Psychology*. Vol.51, NO.6, PP. 309-317.

11. Yang He.(2017).” Robust Inference in Fuzzy Regression Discontinuity Designs”. Computer Science, Mathematics.

12. Yang He and Otávio Bartalotti.(2020).”Wild bootstrap for fuzzy regression discontinuity designs: obtaining robust bias-corrected confidence intervals”. The Econometrics Journal. Vol.23, N0.2, PP. 211-231.

تقدير التأثير السببي للعلاج باستعمال تصاميم الانحدار الضبابي المنقطع

الباحث/ سجاد سمير عبدالرزاق
قسم الاحصاء - كلية الادارة والاقتصاد
جامعة بغداد-العراق
Sajad.samir1201a@coadec.uobaghdad.edu.iq

أ.د. محمد جاسم محمد
قسم الاحصاء - كلية الادارة والاقتصاد
جامعة بغداد-العراق
m.jasim@coadec.uobaghdad.edu.iq

Received:13/11/2022 Accepted:4/12/2022 Published: December / 2022

مستخلص البحث:

في بعض الحالات يحتاج الباحثيين الى معرفة التأثير السببي للعلاج وذلك لمعرفة مدى تأثير العلاج على العينة من اجل الاستمرار باعطاء العلاج او ايقاف العلاج لعدم الفائدة منه. تم استعمال طريقة المربعات الصغرى الموزونة في تقدير معالم الانموذج واستعمال طريقة متعددة الحدود المحلية لتقدير عرض الحزمة. وتم اجراء تجربة في معد الابتكار لدروس التقوية في عام 2021 ل 72 طالبا مشاركا في المعهد. وتم ايجاد مقدر التأثير السببي للعلاج لمجموعة من طلاب معهد الابتكار لدروس التقوية باستعمال الانحدار المنقطع الضبابي ، ووجد ان الاشخاص الذين استخدموا العلاج حصلت زيادة في درجاتهم بعد العلاج بنسبة 25.95% .

نوع البحث: ورقة بحثية.

المصطلحات الرئيسة للبحث: الانحدار المنقطع الضبابي ، التأثير السببي للعلاج ، متعددة الحدود المحلية الحصينة ، طريقة تغطية متوسط مربعات الخطا.

*البحث مستل من رسالة ماجستير