ECO3211 Problem Set 4
Regression Discontinuity Design
Use the 2010 China Family Panel Studies to evaluate the compulsory schooling policy in China.
Table 1. Timing of compulsory schooling reforms and corresponding cutoff points

Province/ Municipality	Effective date	Affected cohort born from
Beijing	6-Aug-86	1-Sep-70
Tianjin	12-Nov-86	1-Sep-70
Hebei	1-Jul-86	1-Sep-70
Shanxi	1-Jul-86	1-Sep-70
Inner Mongolia	15-Sep-88	1-Sep-72
Liaoning	1-Jul-86	1-Sep-70
Jilin	20-Feb-87	1-Sep-70
Helongjiang	1-Jul-86	1-Sep-70
Shanghai	9-Jan-85	1-Sep-69
Jiangsu	9-Sep-86	1-Sep-70
Zhejiang	1-Sep-85	1-Sep-69
Anhui	1-Sep-87	1-Sep-71
Fujian	1-Aug-88	1-Sep-72
Jiangxi	1-Feb-86	1-Sep-69
Shandong	12-Sep-86	1-Sep-70
Henan	1-Oct-86	1-Sep-70
Hubei	1-Mar-87	1-Sep-70
Hunan	1-Sep-91	1-Sep-75
Guangdong	7-Oct-86	1-Sep-70
Guangxi	1-Sep-91	1-Sep-75
Hainan	16-Dec-91	1-Sep-75
Chongqing	1-Jul-86	1-Sep-70
Sichuan	14-Jan-86	1-Sep-69
Guizhou	1-Jan-88	1-Sep-71
Yunnan	29-Oct-86	1-Sep-70
Tibet	1-Jul-94	1-Sep-78
Shaanxi	1-Sep-87	1-Sep-71
Gansu	3-Sep-90	1-Sep-74
Qinghai	1-Oct-88	1-Sep-72
Ningxia	1-Jul-86	1-Sep-70
Xinjiang	28-May-88	1-Sep-71

Steps:

1. Drop the top 10% and bottom 10% of birth cohorts.

2. Define the forcing variable (x), and then normalize it to the cutoff point. Define the location dummy d and generate interaction terms between d and x.

3. Validity test of RDD using the nonparametric approach (Please provide both graphical evidence and regression results)

A. Provide some anecdotal evidence to argue that parents who gave birth to children around the cutoff points could not fully manipulate the forcing variable (year month of birth: x).

B. Provide a density test of x.

C. Check the smoothness of predetermined characteristics: gender, had urban hukou at the age of 3, had urban hukou at the age of 12, being Han group, father's total years of education, and mother's total years of education.

4. Use the nonparametric approach to check the impact of compulsory education on total years of schooling (Please provide both graphical evidence regression results), interpret your findings.

5. Check the robustness of your findings by varying the bandwidth choices and using the parametric estimation approach.

Notes: Assignment 4 will due on **Dec. 13th, 2020**. You should upload three separate documents electronically: one that contains your typed answers to the problem set questions, one well-organized and well-commented Stata do file, and one Stata log file.