Long Question 5 - Flowers in China

Bookmark this page

China suffers from enormous gender imbalance: there are many more boys than girls. Part of this is due to selective abortion, and part is due to worse treatment of girls.

Nancy Qian was interested in finding out whether parents consider the future possible wages of a girl when deciding how much to feed them and take care of them. To this end, she exploits the reform that brought household responsibility system reform in China. After Den Xiao Ping replaced Mao in 1979, households were given the choice about what crop to grow (before, they essentially had to grow cereals), and suitable regions started producing tea and orchards.

Women are particularly useful for tea growing, which requires nimble hands. Therefore she proposed that parents would start taking better care of girls in regions that produce tea. Girls would be more likely to survive, and this would translate into a relatively lower share of males in those region after the reform, thus justifying a difference in difference approach.

Let $POST$ be a dummy for post reform, and TEA be a dummy for whether the region produces tea. Let Y_{it} be the fraction of boys in region i at time t

She runs the following regression:

$$Y_{it} = \beta_0 + \beta_1 TEA_i + \beta_2 POST_t + \beta_3 POST_t * TEA_i + \epsilon_{it}$$

Question 17

0.666666666667/1.0 point (graded)

This question has 3 parts:

Given the regression she runs, which of the following denotes the the average fraction of males in tea-regions, pre-reform?

- β_0
- β_1
- β_2
- β_3
- $\beta_0 + \beta_1$ ✔
- $\beta_1 + \beta_2$
- $\beta_2 + \beta_3$

Given the regression she runs, which of the following denotes the the average fraction of males in non-tea regions, pre-reform?

- (selected) β_0 ✔
- β_1
- β_2
- β_3
- $\beta_0 + \beta_1$
- $\beta_1 + \beta_2$
- $\beta_2 + \beta_3$

In this strategy, which coefficient gives her the causal effect of growing tea on the average fraction of males?

- β_0

- β_1
- β_2
- β_3
- $\beta_0 + \beta_1$
- $\beta_1 + \beta_2$
- $\beta_2 + \beta_3$ ✖

Explanation

Part I. $\mathbb{E}[Y_{it}|TEA_i = 1, POST = 0] = \beta_0 + \beta_1$

Part II. $\mathbb{E}[Y_{it}|TEA_i = 0, POST = 0] = \beta_0$

Part III. $\textbf{causal effect} = [\mathbb{E}[Y_{it}|TEA_i = 1, POST = 1] - \mathbb{E}[Y_{it}|TEA_i = 1, POST = 0] - [\mathbb{E}[Y_{it}|TEA_i = 0, POST = 1] - \mathbb{E}[Y_{it}|TEA_i = 0, POST = 0]]$
$= [\beta_0 + \beta_1 + \beta_2 + \beta_3 - \beta_0 - \beta_1] - [\beta_0 + \beta_2 - \beta_0]$
$= \beta_2 + \beta_3 - \beta_2$
$= \beta_3$

Submit You have used 1 of 1 attempt

Question 18

0.0/1.0 point (graded)

True or False? Instead of including the TEA" dummy, she could include one dummy for each of the regions (excluding one).

- True
- False ✖

Explanation

Including the TEA dummy controls for inherent differences between regions that grow tea and regions that don't grow tea. However, including region fixed effects controls for inherent differences between regions.

Submit You have used 1 of 1 attempt

Question 19

0.0/1.0 point (graded)

This question has 2 parts.

Please round your answer to 2 decimal points

Table 4 – Differences-in-Differences Estimates
of the Effect of Planting Tea and Orchards on Sex Ratios:
Coefficients of the Interactions between Dummies Indicating Whether a Cohort was Born Post Reform
and Dummies Indicating Whether Any Tea Was Planted in the County of Birth

	Dependent Variable : Fraction of Male			
	(1)	(2)	(3)	(4)
Tea * Post	-0.0081	-0.0086	-0.0074	-0.0074
	(0.0024)	(0.0026)	(0.0026)	(0.0026)
Orchard * Post			0.0096	0.0093
			(0.0033)	(0.0033)
Cashcrop * Post		0.0007	-0.0016	-0.0016
		(0.0007)	(0.0011)	(0.0011)
Han	N	N	N	Y
Observations	49082	49082	49082	49082
R-squared	0.09	0.09	0.09	0.09

All regressions include county fixed effect and controls for post and cash crops *post.
Orchard and cashcrop are dummy variables for the amount of orchards and cashcrop planted in each county.
Post = 1 for cohorts born 1979-1990.
Standard errors clustered at county level.

I. Look at column 1 in the table above, what is the t-statistic for the hypothesis H_0 that the coefficient on tea*post is zero?

-3.375 ✖ **Answer:** -3.38

-3.375

II. What is the 90 confidence interval for the coefficient tea*post. Enter the lower and upper bounds on the interval $[a, b]$.

Please round your answer to 2 decimal points

a :

0.14 ✖ **Answer:** -0.01

0.14

b :

0.14 ✖ **Answer:** -0.00

0.14

Explanation

The t-statistic is the point estimate divided by its standard error. In this case, the standard error is in shown in parenthesis.

$$t - statistic = -0.0081/0.0024$$

The CI is: $a = -0.0081 - 1.645 * 0.0024 = -0.012048$
$b = -0.0081 + 1.645 * 0.0024 = -.004152$

Submit You have used 1 of 1 attempt

Question 20

1.0/1.0 point (graded)

We only observe tea production in the regions that have *chosen* to produce tea. Your friend who is an anthropology major argues that in some regions, people are more likely to prefer girls, for historical reasons. Could these regions then decide to grow more tea?

If yes, what assumption underlying this design strategy would this violate?

- the exclusion restriction

- (•) the parallel trends assumption ✔
- ( ) the independence assumption
- ( ) None of the above

Explanation
If regions that grew tea and regions that did not grow tea, would follow different trends at the time of the policy change, then even they maintained parallel trends before the reform. The effect of tea growth would be conflated with this inherent difference.

Submit You have used 1 of 1 attempt

Question 21

0.0/1.0 point (graded)
To solve this problem, Qian uses the fact that some regions are more suitable to tea production than others: in particular, a certain amount of rain, elevation and slopes are needed to produce tea. She decides to propose an instrumental variables strategy.

What is the firststage equation?
(Select one)

- An OLS regression of TEA on rain, elevation, and slope.
- An OLS regression of the fraction of boys on rain, elevation, and slope.
- An OLS regression of TEA on rain, elevation, and slope including region and time fixed effects. ✖
- An OLS regression of the fraction of boys on rain, elevation, and slope including region fixed effects and a dummy for whether or not the country grows Orchards.

Explanation

Qian's proposed IV strategy is to use geographic features (rain, elevation, and slope) as an instrument for whether or not a region grows tea. So the first stage is an OLS regression of TEA(the variable that is instrumented for) on the instruments.

Submit You have used 1 of 1 attempt

Question 22

1.0/1.0 point (graded)

What assumptions are needed for the instrument to be a good instrument?

(Select one)

- Rain, elevation, and slope vary randomly across regions that grow tea and regions that don't.
- Rain, elevation, and slope affect whether or not a region is suitable for tea growth.
- Rain, elevation, and slope don't affect the fraction of boys except through tea growth.
- All of the above. ✔

Explanation
Refer back to the lecture questions for the past 2 lectures- all of the above are necessary conditions for the instrument's validity.

Submit You have used 1 of 1 attempt