Econometrics HW2

Michael B. Nattinger[*]

March 25, 2021

1 Question 18.2

1.1 Part A

This is essentially just a regression of Y on D with state fixed effects, so the estimator is the following:

$$\hat{\theta} = \frac{\sum_{t=0}^{T} \sum_{i=0}^{N} (D_{it} - \bar{D}_i)(Y_{it} - \bar{Y}_i)}{\sum_{t=i}^{T} \sum_{i=1}^{N} (D_{it} - \bar{D}_i)^2},$$

where $\bar{D}_i = (1/n) \sum_{t=1}^{n} D_{it}, \bar{Y}_i = (1/n) \sum_{t=1}^{n} Y_{it}$.

1.2 Part B

For the untreated sample, $D_{it} = 0 \forall t$, so $D_{it} - \bar{D}_i = 0 \forall t$, so $\hat{\theta} = \frac{\sum_{t=0}^{T} \sum_{i=0}^{N} (D_{it} - \bar{D}_i)(Y_{it} - \bar{Y}_i)}{\sum_{t=0}^{T} \sum_{i=0}^{N} (D_{it} - \bar{D}_i)^2} = \frac{\sum_{t=0}^{T} (D_{1t} - \bar{D}_1)(Y_{1t} - \bar{Y}_1)}{\sum_{t=0}^{T} (D_{1t} - \bar{D}_1)^2}$ which is a function only of the treated sample.

1.3 Part C

No, it is only a difference estimator of the treated group. It is not accounting for the control group, as shown from the independence result from Part B.

1.4 Part D

If the time trend is not important then the control group would have no change over time, so the difference estimator would be the same as the difference-in-difference estimator.

2 Question 18.4

For both, if all interaction dummies were included then there would be perfect collinearity and the $X'X$ matrix would not be invertible. Intuitively one of the groups serve as a baseline from which the other groups are compared.

3 Question 18.5

3.1 Part A

The difference in Wisconsin is $16.72 - 15.23 = 1.49$. The difference in Minnesota is $18.01 - 16.42 = 1.59$. The difference in difference estimate is therefore $1.49 - 1.59 = -0.1$.

[*] I worked on this assignment with my study group: Alex von Hafften, Andrew Smith, and Ryan Mather. I have also discussed problem(s) with Emily Case, Sarah Bass, Katherine Kwok, and Danny Edgel.

3.2 Part B

$\hat{\beta} = -0.1.$

3.3 Part C

$\hat{\gamma} = 15.23 - 16.42 = -1.19.$

4 Question 17.1

4.1 Part A

We will integrate and make a variable substitution $y = \frac{X_i - x}{h} \Rightarrow hdy = dx$

$$\begin{aligned} E[X^*] &= \int x \frac{1}{nh} \sum_{i=1}^{n} K((X_i - x)/h) dx \\ &= \frac{1}{nh} \sum_{i=1}^{n} \int x K((X_i - x)/h) dx \\ &= \frac{1}{n} \sum_{i=1}^{n} \int (X_i + hy) K(y) dy \\ &= \frac{1}{n} \sum_{i=1}^{n} X_i \int K(y) dy + \frac{h}{n} \sum_{i=1}^{n} \int y K(y) dy \\ &= \frac{1}{n} \sum_{i=1}^{n} X_i \\ &= \bar{X}_n, \end{aligned}$$

where we have used the fact that the kernels integrate to 1 and have mean 0.

4.2 Part B

We will again ust the same substitution.

$$\begin{aligned} Var(X^*) &= E[X^{*2}] - E[X^*]^2 \\ &= \int x^2 \frac{1}{nh} \sum_{i=1}^{n} K((X_i - x)/h) dx - \bar{X}_n^2 \\ &= \frac{1}{n} \sum_{i=1}^{N} \int (X_i + hy)^2 K(y) dy - \bar{X}_n^2 \\ &= \frac{1}{n} \sum_{i=1}^{N} X_i^2 \int K(y) dy + \frac{2h}{n} \sum_{i=1}^{N} X_i \int y K(y) dy + \frac{h^2}{n} \sum_{i=1}^{N} \int y^2 K(y) dy - \bar{X}_n^2 \\ &= \frac{1}{n} \sum_{i=1}^{N} X_i^2 + h^2 - \bar{X}_n^2 \\ &= \hat{\sigma}^2 + h^2. \end{aligned}$$

5 Question 17.3

The optimal bandwidth is $h_0 = \left(\frac{R_K}{R(f'')}\right)^{1/5} n^{-1/5}$. Note that, for a uniform distribution, $f'' = 0 \Rightarrow R(f'') = 0$. The optimal bandwidth is the largest feasible bandwidth.

6 Question 17.4

The effective bandwidth would be much larger as the scale was reduced by a factor of 10^6 yet the bandwidth was unchanged. The plot would be much wider, smoother, and lower than it should be.

If the bandwidth was scaled appropriately by 1000000^{-1} when the scale changed, the density plots would have the same shape.

7 Question 19.3

When $m(x)$ is increasing and convex, the bias is positive. When $m(x)$ is increasing and concave, the bias is negative. When $m(x)$ is decreasing and concave, the bias is again negative. When $m(x)$ is decreasing and convex, the bias is positive. The intuition here is that you are locally averaging around a point of interest. This is similar to taking the midpoint of a function over a region of concavity or convexity.

Moreover, asymptotic theory tells us that the bias is $\frac{1}{2}m''(x)h^2$ and $m''(x) > 0$ for a convex function and $m''(x) < 0$ for a concave function.

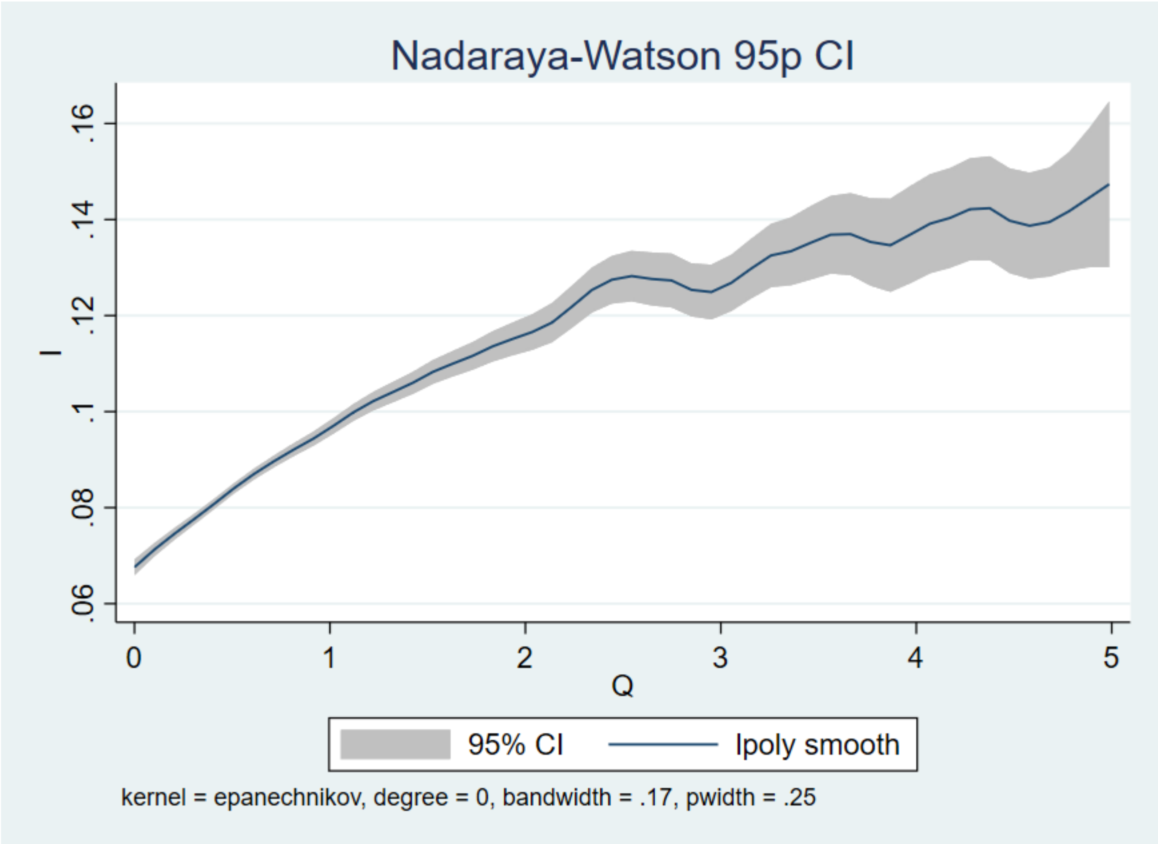
8 Question 19.4

Bias is $B(x) = (1/2)m''(x) + m'(x)f'(x)/f(x) = \beta f'(x)/f(x)$. If $\beta > 0$, then $b(x) < 0 \iff f'(x) < 0$, $b(x) > 0 \iff f'(x) > 0$. In contrast, if $\beta < 0$ then $b(x) < 0 \iff f'(x) > 0$, $b(x) > 0 \iff f'(x) < 0$.

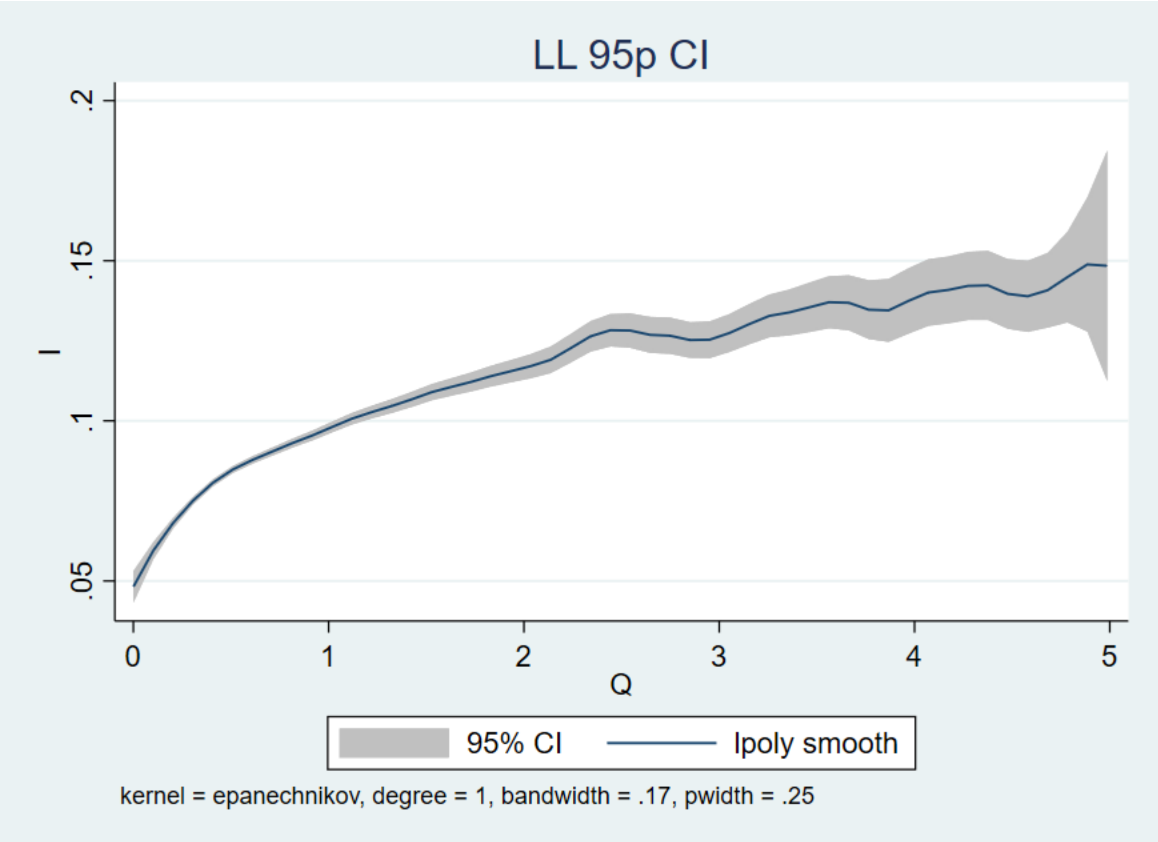
$f(x)$ is the marginal density of X. If $f'(x) > 0$ then the density of X is increasing at x, so if β is positive (negative), the bias at x is positive because it is influenced more by the mass of X to the right of x, so the positive (negative) influence of this mass on the conditional expectation due to the positive (negative) value of β leads to a positive (negative) bias.

9 Question 19.9

9.1 Part A



9.2 Part B



9.3 Part C

Yes, from the graphs in the preceding parts of this question it appears that there is nonlinearity in the relationship.

Stata code:

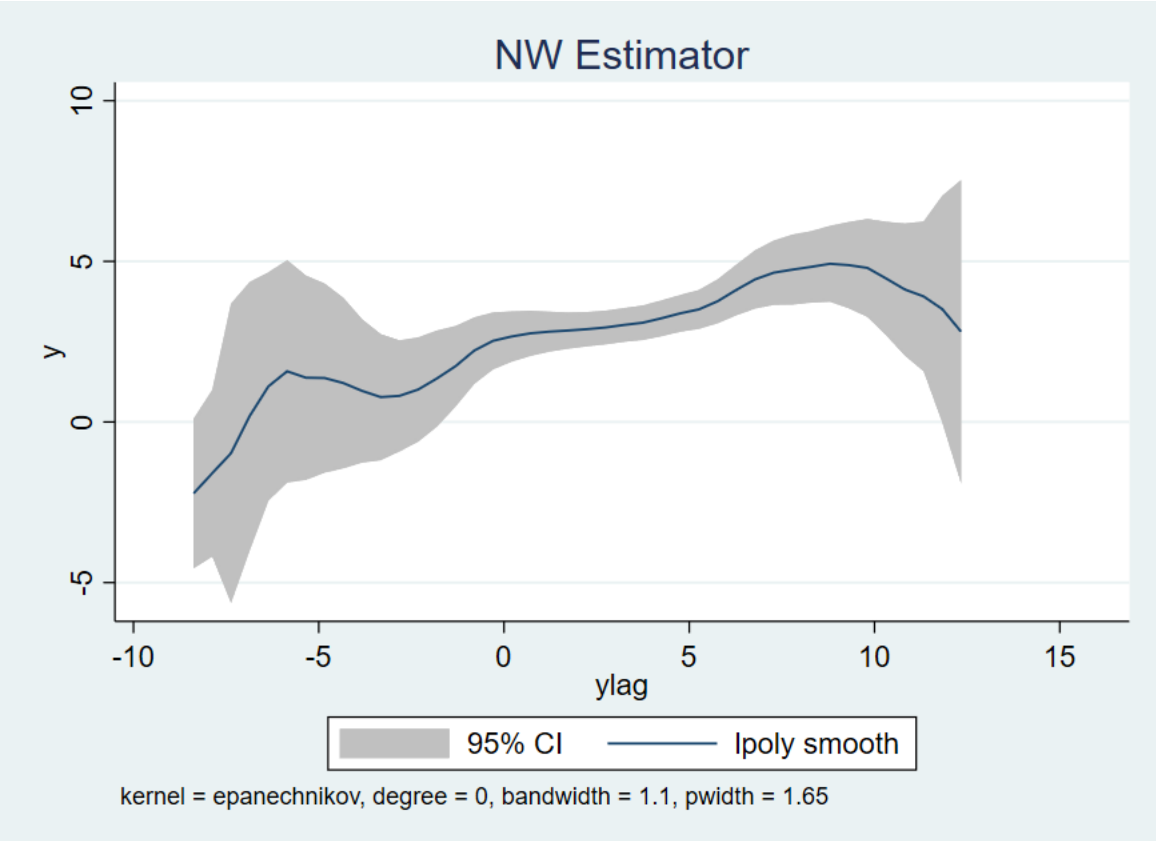
```
pause on
use "C:\Users\micha\OneDrive\Documents\HOMEWORK\Y1S1\Metrics\Q4\PS2\Invest1993", clear

gen I = inva
gen Q = vala
lpoly I Q if Q<=5, ci nosc title("Nadaraya-Watson 95p CI")
graph export "C:\Users\micha\OneDrive\Documents\HOMEWORK\Y1S1\Metrics\Q4\PS2\pings\9p1.png", as(
png) name("Graph")
lpoly I Q if Q<=5, degree(1) ci nosc title("LL 95p CI")
graph export "C:\Users\micha\OneDrive\Documents\HOMEWORK\Y1S1\Metrics\Q4\PS2\pings\9p2.png", as(
png) name("Graph")
```

10 Question 19.11

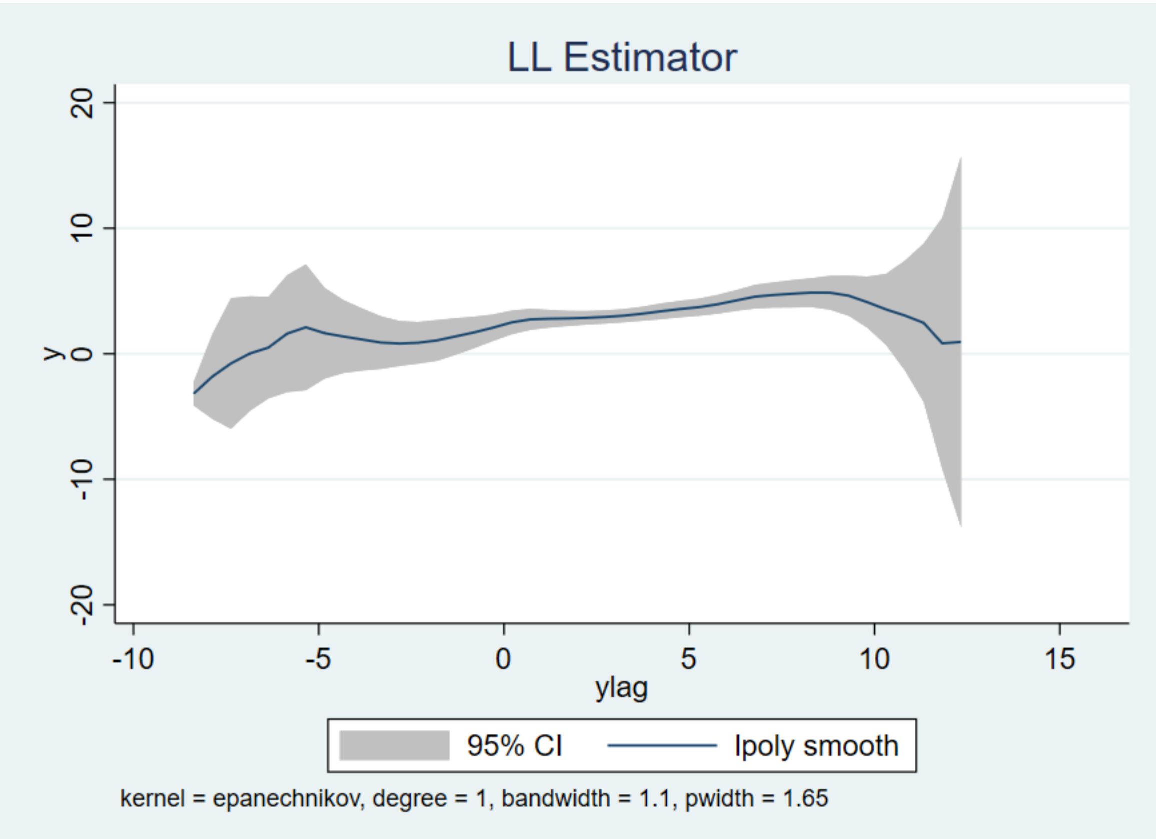
10.1 Part A

Done in stata, see code below the final part of this question.

10.2 Part B



10.3 Part C



10.4 Part D

Yes, from the graphs in the preceding parts of this question it appears that there is nonlinearity in the relationship.

Stata code:

```
pause on
use "C:\Users\micha\OneDrive\Documents\HOMEWORK\Y1S1\Metrics\Q4\PS2\FRED-QD", clear

gen y = 100*((gdpc1/L.gdpc1)^4-1)
gen ylag = L.y
lpoly y ylag, ci nosc title("NW Estimator")
graph export "C:\Users\micha\OneDrive\Documents\HOMEWORK\Y1S1\Metrics\Q4\PS2\pings\11p1.png",
as(png) name("Graph")
lpoly y ylag, degree(1) ci nosc title("LL Estimator")
graph export "C:\Users\micha\OneDrive\Documents\HOMEWORK\Y1S1\Metrics\Q4\PS2\pings\11p2.png",
as(png) name("Graph")
```