Problem Set #9

Danny Edgel
Econ 710: Economic Statistics and Econometrics II
Spring 2021

April 2, 2021

Discussed and/or compared answers with Sarah Bass, Emily Case, Katherine Kwok, Michael Nattinger, and Alex Von Hafften

Exercise 20.1

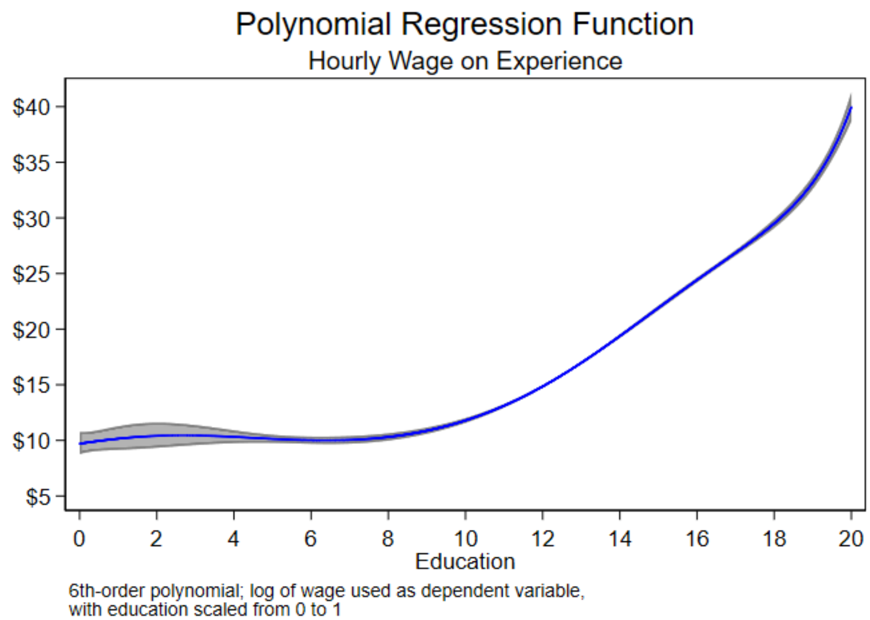
If $X = 3$, then all three dummies for X are equal to 1, so the marginal effect is $2 + 5 - 3 = 4$.

Exercise 20.3

$m_K(x)$ is concave if $\beta_1 > 0$ and $\beta_j < 0$ for $j \in \{2, 3, 4\}$.

Exercise 20.11

The graph below displays the results of the polynomial regression, with both experience and wage rescaled to their absolute levels.



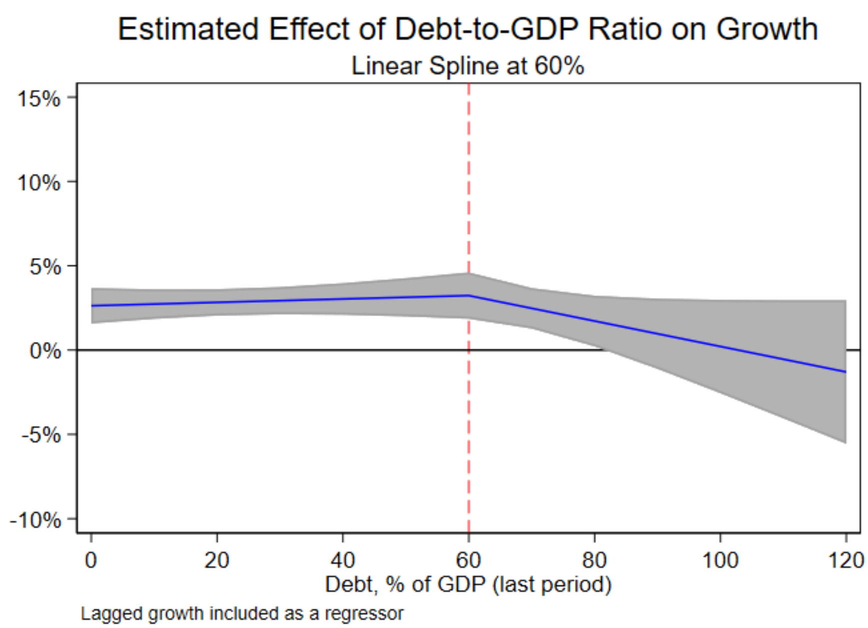
Exercise 20.15

The results of the three regression estimations and the plot of $\hat{m}(x)$ for the second model are below. The AIC for each model is displayed at the bottom of the regression table. The AIC suggests that the linear model is half as probable as the two spline models, which are equally likely. This suggests that the true relationship between the debt-to-GDP ratio may have a slope discontinuity, but it is not clear from the analysis shown here.

VARIABLES	(1) Y_t	(2) Y_t	(3) Y_t
Y_{t-1}	0.300***	0.284***	0.280***
	(0.0654)	(0.0658)	(0.0663)
D_{t-1}	-0.00802	0.0101	0.0343
	(0.0117)	(0.0156)	(0.0241)
$(D_{t-1} - 40)\mathbb{1}\{D_{t-1} \geq 40\}$			-0.0871
			(0.0560)
$(D_{t-1} - 80)\mathbb{1}\{D_{t-1} \geq 80\}$			-0.0183
			(0.104)
$(D_{t-1} - 60)\mathbb{1}\{D_{t-1} \geq 60\}$		-0.0856*	
		(0.0492)	
Constant	2.898***	2.628***	2.312***
	(0.515)	(0.536)	(0.592)
Observations	218	218	218
R-squared	0.093	0.106	0.114
AIC	1250	1249	1249

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1



Exercise 21.1

When the treatment occurs after the discontinuity, we estimate average treatment effect $\hat{\theta} = \hat{m}_1(c) - \hat{m}_0(c)$. The only difference that arises when the treatment occurs prior to the discontinuity is that $\hat{\theta}$ is now the negative treatment effect.

Exercise 21.2

By the logic of RDD, we have two points over which we can estimate an average treatment effect. Our model is:

$$\mathbb{E}\left[Y|X = x, D\right] = m(x) = m_0(x)\mathbb{1}\left\{X < c_1|X > c_2\right\} + m_1(x)\mathbb{1}\left\{c_1 \leq X \leq c_2\right\}$$

Thus, we can identify either $ATE(c_1) = m_1(c_1) - m_0(c_1)$ or $ATE(c_1) = m_1(c_2) - m_0(c_2)$

Exercise 21.3

Suppose that we have the following model:

$$Y = m(X) + e$$

Where $\mathbb{E}\left[e|X\right] = 0$ and there exists a policy that affects Y and applies whenever $X \geq c$. Then, taking the expectation of each side of the model, we get:

$$\begin{aligned}
\mathbb{E}\left[Y|X = x\right] &= \mathbb{E}\left[Y_0|X = x\right]\mathbb{1}\left\{X < c\right\} + \mathbb{E}\left[Y_1|X = x\right]\mathbb{1}\left\{X \geq x\right\} \\
\mathbb{E}\left[Y_0|X = x\right] &= m_0(x) \\
\mathbb{E}\left[Y_1|X = x\right] &= m_1(x) \\
\Rightarrow m(x) &= m_0(x) = m_0(x)\mathbb{1}\left\{x < c\right\} + m_1(x)\mathbb{1}\left\{x \geq c\right\}
\end{aligned}$$

Exercise 21.4

If we pick a rectangular kernel, then we have kernel function:

$$K(u) = \mathbb{1}\left\{|u| \leq \frac{1}{2}\right\}$$

Thus, the objective function for a local linear estimation with rectangular bandwidth is:

$$\min \sum_{i=1}^{n}(Y_i - \beta_0 - \beta_1 X_i - \beta_2(x_i - c)T_i - \theta T_i)^2\mathbb{1}\left\{|\frac{X_i - c}{2h}| \leq \frac{1}{2}\right\}$$

If the subsample $|X - C| \leq h$ is used, then the indicator function in the objection function is always satisfied, so the first-order condition of the local linear estimation is:

$$Y_i - \beta_0 - \beta_1 X_i - \beta_2(x_i - c)T_i - \theta T_i = 0 \Rightarrow Y_i = \beta_0 + \beta_1 X_i + \beta_2(x_i - c)T_i + \theta T_i$$

Which is equivalent to equation (21.4).

Exercise 21.6

The results are displayed below and generated in the attached .do file.

	$h = 4$		$h = 12$	
	Baseline	Covariates	Baseline	Covariates
$\hat{\theta}$	-3.049	-3.031	-1.774	-1.760
$s(\hat{\theta})$	(1.333)	(1.329)	(0.844)	(0.838)
% Black		0.012		0.018
$s(\hat{\beta}_1)$		(0.020)		(0.010)
% Urban		-0.018		-0.020
$s(\hat{\beta}_2)$		(0.019)		(0.010)

Exercise 21.8

The table below uses the *mort_age25plus_related_postHS* variable instead of *mort_age59_related_postHS*, as in exercise 21.6.

	$h = 4$		$h = 12$	
	Baseline	Covariates	Baseline	Covariates
$\hat{\theta}$	2.425	2.478	3.705	4.118
$s(\hat{\theta})$	(7.704)	(7.830)	(5.093)	(5.119)
% Black		-0.195		-0.191
$s(\hat{\beta}_1)$		(0.116)		(0.062)
% Urban		0.012		-0.150
$s(\hat{\beta}_2)$		(0.114)		(0.064)