Ec141, Spring 2020

Midterm 2

Please read each question carefully. Start each question on a new sheet or paper. Turn in your exam on bCourses as a pdf file (you may find the iOS app "Scannable" or something similar useful for making your pdf file). Please make sure you scan is readable and clean, the pages correctly ordered in your file and so on. Allow time to prepare you file for submission prior to the exam due date/time of 11AM Saturday May 2nd.

This exam is open book. You are not allowed to consult with others however – this includes fellow classmates or other individuals who might be knowledgeable about the material. You are free to consult your course notes, assigned readings, etc.

Good luck!

[1] **[5 Points]** Please write your full name and student ID on the first page of your pdf file of exam solutions.

[2] **[30 Points]** You've been hired by the Government of Honduras to assess the efficacy of treatment for decompression sickness among lobster divers in La Moskitia. In this region of Honduras lobsters are harvested by divers who, on occasion, get decompression sickness which may result in partial paralysis or worse. You are provided the following table of information about 300 diving accident victims.

		$Y=0$ (No Limp)	$Y=1$ (Limp)
$X=0$ (Untreated)	$W=0$ (Depth $<$ 75')	90	10
	$W=1$ (Depth $\geq$ 75')	10	40
$X=1$ (Treated)	$W=0$ (Depth $<$ 75')	30	20
	$W=1$ (Depth $\geq$ 75')	50	50

[a] **[2 Points]** What is the probability of a victim walking with a limp conditional on treatment ($X=1$) and non-treatment ($X=0$)?

[b] **[3 Points]** What is the probability of a victim receiving treatment conditional on having dived "deep" ($W=1$) vs. "shallow" ($W=0$)?

[c] **[7 Points]** A government official worries that treatment is harming the divers and thinks it would be better to do nothing. Present a counter-argument to this official. **[4 to 6 sentences]**

[d] **[8 Points]** Let $Y(0)$ and $Y(1)$ denote a divers potential outcome given non-treatment and treatment respectively. Discuss the conditional independence assumption assumption

$$(Y(0), Y(1)) \perp X | W = 0, 1.$$

Make a positive and negative argument for this assumption. **[4 to 6 sentences]**

[e] **[5 Points]** Using the assumption in part [d] construct the IPW estimate of the average treatment effect (ATE) on the outcome. Report your result to the government official. Your report should include an explanation for why and how your are adjusting for accident depth. Is treatment effective? **[4 to 6 sentences]**

[f] **[5 Points]** Say instead you were given the table:

		$Y=0$ (No Limp)	$Y=1$ (Limp)
$X=0$ (Untreated)	$W=0$ (Depth <75')	90	10
	$W=1$ (Depth ≥ 75')	0	0
$X=1$ (Treated)	$W=0$ (Depth <75')	30	20
	$W=1$ (Depth ≥ 75')	75	75

Can you compute the ATE is this case? Why or why not? **[2 to 3 sentences]**

[3] **[25 Points]** Let Y equal tons of banana's harvested in a given season for a randomly sampled Honduran banana planation. Output is produced using labor and land according to $Y = AL^{\alpha_0}D^{1-\alpha_0}$, where L is the number of employed workers and D is the size of the plantation in acres and we assume that $0 < \alpha_0 < 1$. The price of a unit of output is P, while that of a unit of labor is W. These prices may vary across plantations (e.g., due to transportation costs, labor market segmentation etc.). We will treat D as a fixed factor; A captures sources of plantation-level differences in farm productivity due to unobserved differences in, for example, soil quality and managerial capacity. Plantation owners choose the level of employed labor to maximize profits. The observed values of L are therefore solutions to the optimization problem:

$$L = \arg\max_{l} P \cdot Al^{\alpha_0}D^{1-\alpha_0} - W \cdot l.$$

[a] **[2 Points]** Show that the amount of employed labor is given by

$$L = \left\{\alpha_0 \frac{P}{W} A\right\}^{\frac{1}{1-\alpha_0}} D. \tag{1}$$

[b] **[3 Points]** Let $a_0 = \frac{1}{1-\alpha_0}\ln\alpha_0 + \frac{1}{1-\alpha_0}\mathbb{E}[\ln A]$, $b_0 = \frac{1}{1-\alpha_0}$, and $V = \frac{1}{1-\alpha_0}\{\ln A - \mathbb{E}[\ln A]\}$. Show that the log of the labor-land ratio is given by

$$\ln\left(\frac{L}{D}\right) = a_0 + b_0 \ln\left(\frac{P}{W}\right) + V \tag{2}$$

and that, letting $c_0 = \mathbb{E}[\ln A]$ and $U = \ln A - \mathbb{E}[\ln A]$, the log of planation yield (output per unit of land) is given by

$$\ln\left(\frac{Y}{D}\right) = c_0 + \alpha_0 \ln\left(\frac{L}{D}\right) + U. \tag{3}$$

[c] **[5 Points]** Briefly discuss the content and plausibility of the restriction **[4 to 6 sentences]**

$$\mathbb{E}[\ln A | \ln(P/W)] = \mathbb{E}[\ln A]. \tag{4}$$

[d] **[5 Points]** Using (2), (3) and (4) show that the coefficient on $\ln(L/D)$ in $\mathbb{E}^*[\ln(Y/D)|\ln(L/D)]$ equals

$$\alpha_0 + (1-\alpha_0)\frac{\mathbb{V}(\ln A)}{\mathbb{V}(\ln A) + \mathbb{V}(\ln(P/W))}.$$

Provide some economic intuition for this result. **[4 to 6 sentences]**

[e] **[5 Points]** Using (2), (3) and (4) show that the coefficient on $\ln(L/D)$ in $\mathbb{E}^*[\ln(Y/D)|\ln(L/D), V]$ equals α_0. Provide some economic intuition for this result. **[4 to 6 sentences]**

[f] **[5 Points]** Assume that all plantations face the same output price (P) and labor cost (W). What value does the coefficient on $\ln(L/D)$ in $\mathbb{E}^*[\ln(Y/D)|\ln(L/D)]$ equal now? Why? **[4 to 6 sentences]**

[4] **[30 Points]** Consider a population of high school graduates. Let Y_1 denote the earnings an individual in this population would get if they completed at least four years of college, let Y_0 denote the earnings they would get if they did not complete college. Let $D = 1$ in an individual actually completes college and zero otherwise. Let Y denote observed earnings which, given the data structure outlined above, equals

$$Y = (1 - D) Y_0 + DY_1.$$

Assume that

$$(Y_1, Y_0) \perp D|X$$

where X is a characteristic measured at the completion of high school but prior to any college attendance. Further assume that

$$\begin{aligned} \mathbb{E}[Y_1|X] &= \alpha_1 + \gamma_1 X \\ \mathbb{E}[Y_0|X] &= \alpha_0 + \gamma_0 X. \end{aligned}$$

Note for the purposes of this problem you may assume that all individuals who attend college graduate and all individuals who do not attend college are high school graduates.

[a] **[7 Points]** Across subpopulations homogenous in $X = x$ can we use Y_1 or Y_0 to predict college attendance? Would these variables be informative about college attendance unconditional on X? Under what conditions can we use an individual's *observed* wage Y to predict whether they went to college? **[4 to 6 sentences]**

[b] **[8 Points]** Let $\beta_0 = E[Y_1 - Y_0|D = 1]$. Interpret this object. **[2 to 3 sentences]** Show that

$$\beta_0 = (\alpha_1 - \alpha_0) + (\gamma_1 - \gamma_0)\mathbb{E}[X|D = 1].$$

Explain the role of the assumptions introduced above in your derivation. **[2 to 3 sentences]**

. [c] **[5 Points]** Show that

$$\mathbb{E}[Y|X, D] = \alpha_0 + \gamma_0 X + (\alpha_1 - \alpha_0) D + (\gamma_1 - \gamma_0) DX.$$

[d] **[10 Points]** Consider the following least squares fit of log earnings on a dummy for completion of an undergraduate degree using a sample of 1,754 white males from the NLSY79.

$$\text{LogEarnings} = \underset{(0.0220)}{10.0332} + \underset{(0.0357)}{0.4879} \quad \text{UNDERGRAD} \quad . \tag{5}$$

Now consider the least squares fit which additionally includes an individual's AFQT percentile score and its

interaction with UNDERGRAD:

$$\begin{aligned}\text{LogEarnings} = & \underset{(0.0573)}{9.8231} + \underset{(0.0010)}{0.0040} \quad \text{AFQT} \\ & + \underset{(0.1584)}{0.1898} \quad \text{UNDERGRAD} + \underset{(0.0020)}{0.0023} \quad \text{UNDERGRAD} \times \text{AFQT}\end{aligned} \tag{6}$$

Finally consider the least squares fit of AFQT on a constant and UNDERGRAD:

$$\text{AFQT} = \underset{(0.73)}{52.27} + \underset{(1.05)}{28.35} \quad \text{UNDERGRAD} \quad . \tag{7}$$

Assume that $X = AFQT$. Using these results compute an estimate of β_0. Justify and explain your calculations. **[2 to 3 sentences]** How does your estimate compare with the coefficient on UNDERGRAD in (5)? Comment on an differences and provide an explanation for them (if needed). **[4 to 6 sentences]**

[5] **[10 Points]** The Vice Chancellor for Undergraduate Education is concerned about students dropping out for Cal prior to finishing the requirements for a BA. She provides you with the following Table. The table refers to the Cal students who first arrived on campus in the Fall semester of 2013.

	Number in F13 Still at Cal	Number Dropping out	Number Transferring	Hazard	Survival	Std. Error
F13	6,000	500	200			
S14		530	70			
F14		940	260			
S15		350	150			

The "Number Transferring" column reports the number of students who transfer to another University at the close of the semester. You may assume that these students are lost to further follow-up. The "Std. Error" column refers to the standard error of the survival function.

[a] **[5 Points]** State and discuss the "random censoring" assumption introduced in lecture. Is this assumption credible in the current context? Explain. **[4 to 6 sentences]**

[b] **[5 Points]** Under the maintained assumption of random censoring fill-in the empty cells in the table. What is the median number of semesters enrolled at Cal prior to drop-out.