Due on Canvas site by midnight of Monday, October 2. You may work together in small groups, but each student must write up his or her answers separately and list the names of the other students with whom he or she worked.

1. The Ministry of Education in Ghana wishes to evaluate two different programs for supporting primary school students who are falling behind in math and language. It wishes to determine whether either program, or both, can improve average learning outcomes for an entire school. The first program (the "Afterschool Program") requires students who are falling behind to remain after school three days a week to take remedial classes. The second program (the "Tracking Program") has no afterschool component, but instead sorts students by ability level within the classes during the regular school day. Students who are falling behind are assigned to separate classes from those not falling behind. The outcome of interest is the average test score of all students at a given school.
    a. Below are three evaluation designs that have been proposed. For each, briefly identify one or more concerns the design raises.
        i. Every government school in the country selects the program that administrators believe will work best for their school. Compare student test scores of schools that selected the Afterschool Program and of those that selected the Tracking Program to test scores of schools who selected neither.

           This sort of selection bias may not yield representative or accurate results as there may be a similarity or relationship between the administration's choice and the performance of their students. There may also be a spillover effect in which maybe administrators in a similar geographic region may all decide on the same program and maybe this area has higher or lower performance compared to another district.

        ii. Evaluators select a random sample of 90 government schools from around Ghana. The administrator of each school indicates her preferences for one or more of the programs, and the evaluators assign 30 schools to the Afterschool Program, 30 to the Tracking Program, and 30 to neither based on these preferences. Compare average test scores of each group of 30 schools.

Though the random sampling of schools improves the methods slightly, the assignment of programs based on administrative preferences still allows for bias to enter and skew the end results. We don't know if or to what extent administrative choice is a covariate or has a relationship with student performance.

iii. Half of the schools in the country are randomly selected to implement the Afterschool Program and the other half to implement the Tracking Program. Compare average tests of schools in each of these groups.

The random selection of government schools and the random assignment of intervention programs makes this a stronger proposition than the previous two, however having a control group would make this proposal even stronger. Instead I'd break down the sample into three groups: 1) Afterschool 2) Tracking and 3) Control.

Due to political and financial constraints, the Ministry decides to pilot only the Afterschool Program and to run a small evaluation for this pilot as well. Five schools from near Accra (the capital of Ghana) volunteer to participate. Three of these schools are randomly selected to implement the Afterschool Program. The data reported below are from one year after these schools implemented the program.

School	Afterschool Program?	Average Test Score (out of 100)
Accra I	Yes	81
Accra II	No	70
Kasoa	No	77
Madina	Yes	78
Taifa	Yes	86

b. What is the estimate of the effect of the Afterschool Program on average test score based on these data?

```
. regress score inter

      Source |       SS           df       MS      Number of obs   =         5
-------------+----------------------------------   F(1, 3)         =      4.20
       Model |  80.0333333         1  80.0333333   Prob > F        =    0.1328
    Residual |  57.1666667         3  19.0555556   R-squared       =    0.5833
-------------+----------------------------------   Adj R-squared   =    0.4444
       Total |       137.2         4        34.3   Root MSE        =    4.3653

------------------------------------------------------------------------------
       score |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       inter |   8.166667   3.984925     2.05   0.133    -4.515144    20.84848
       _cons |       73.5    3.08671    23.81   0.000     63.67671    83.32329
------------------------------------------------------------------------------
```

With this small sample size, the estimate of effect of the Afterschool program on average test school, based on the data provided is +8.167

Schools with the afterschool intervention had 8.167% higher score average than those schools not participating in the afterschool intervention.

This can be calculated by running a regression in Stata or calculating the differences in score averages between those in the intervention and the control schools.

i	Y_{1i}	Y_{0i}	D_i	Outcome Observed Y_i	Causal Effect of Treatment Y_{1i}- Y_{0i}
Accra I	[redacted]	[redacted]	1	81	[redacted]
Accra II	[redacted]	[redacted]	0	70	[redacted]
Kasoa	[redacted]	[redacted]	0	77	[redacted]
Madina	[redacted]	[redacted]	1	78	[redacted]
Taifa	[redacted]	[redacted]	1	86	[redacted]

Naïve estimate of causal effect = E[Y|D=1] – E[Y|D=0] = 81.667 - 73.5 = 8.167

c. Conduct a Fisher's Exact Test using these data. Be sure to precisely specify the null hypothesis and to compute and interpret the P-value.

H_0 / Null Hypothesis: $Y_1 - Y_0 = 0$ **Sharp Null*

H_A / Alternative Hypothesis: $Y_1 - Y_0 \neq 0$ **or > 0**

N = 5 **3 = D =1** **2 = D = 0**

Alpha / α **= 95% / 0.05**

	Accer I	Madina	Taifa	Accera II	Kasoa	$\overline{Y_1} - \overline{Y_0}$
Y_i	81	78	86	70	77	8.167
D_i	1	1	1	0	0	-

P	1	1	1	0	0	8.167
E	1	1	0	1	0	0.167
R	1	1	0	0	1	0.667
M	1	0	1	1	0	1.500
U	1	0	1	0	1	7.333
T	1	0	0	1	1	-6.00
A	0	1	1	1	0	-1.00
T	0	1	0	1	1	-8.50
I	0	1	1	0	1	4.833
O N	0	0	1	1	1	-1.833

S						

Out of the 10-possible permutation of treatment assignments, 2 would yield at least as large an absolute difference | $\overline{Y_1} - \overline{Y_0}$| as the realized permutation, assuming the sharp null was true.

Assuming the sharp null, the probability of observing the difference we observed or something more extreme is 2/10 = 0.20

The p-value of conducting a fisher's exact test in by means of permutations indicates that the observed differences in our sample permutation is not statistically significant.

d. Assume (for this part of the question only) that this evaluation was conducted properly and that it concluded that the Afterschool Program did have a statistically significant effect on test scores (regardless of your answer from part c). Now the Ministry wants to scale this program up to include all schools in Ghana. Give three concerns you have related to the external validity of this finding and the Ministry's desire to scale up.

The small sample size of n=5 is of concern and a threat to external validity. It jeopardizes the generalizability and with so few observations we can't entirely that the dependent variable of interest, scores, is attributable to another covariate. The findings well represent our small sample but even with a statistically significant finding, our confidence in stating that these results will be reproducible on much larger scale is probably not enough to allocate funds to back such an initiative. We also have not tested or assessed the "Tracking Program" and don't know if it would yield even better estimates of causal effect.

2. This question requires you to interpret data reported in Alan Krueger's paper, "Experimental Estimates of Education Production Functions" (1999). a. For each of these six characteristics, explain whether the value of the P-value reported in the right-most column is cause for concern. In your answer, be sure to explain the meaning of the P-value in this context.

Table I compares mean values of six characteristics for students assigned to each of the three treatment arms (small, regular, and regular/aide). A portion of that table is reproduced here:

TABLE I
COMPARISON OF MEAN CHARACTERISTICS OF TREATMENTS AND CONTROLS: UNADJUSTED DATA

A. Students who entered STAR in kindergarten[b]

Variable	Small	Regular	Regular/Aide	Joint *P*-Value[a]
1. Free lunch[c]	.47	.48	.50	.09
2. White/Asian	.68	.67	.66	.26
3. Age in 1985	5.44	5.43	5.42	.32
4. Attrition rate[d]	.49	.52	.53	.02
5. Class size in kindergarten	15.1	22.4	22.8	.00
6. Percentile score in kindergarten	54.7	49.9	50.0	.00

a. For each of these six characteristics, explain whether the value of the P-value reported in the right-most column is cause for concern. In your answer, be sure to explain the meaning of the P-value in this context.

Table I is a balance test of this study and the p-value in Table I is an F-test of equality of all three groups. And it should be noted that these p-values are not conditional on things like school-effects.

Above we're testing / looking at the overall variation over categories to see if students were successfully randomly assigned across class types in the STAR program.

1. Free Lunch
There is a 9% chance of observing a sample such as this one or more extreme, given our null hypothesis is true. This is greater than the standard alpha level of 0.05 meaning we reject the null hypothesis. It also means that it's likely students were sufficiently randomized across class type by student's receiving free lunch in their first school year.

2. White/Asian
There is a 26% chance of observing a sample such as this one or more extreme, given our null hypothesis is true. This is greater than the

standard alpha level of 0.05 meaning we reject the null hypothesis. It also means that it's likely students were sufficiently randomized across class type by ethnicity.

3. Age in 1985
There is a 32% chance of observing a sample such as this one or more extreme, given our null hypothesis is true. This is greater than the standard alpha level of 0.05 meaning we reject the null hypothesis. It also means that it's likely students were sufficiently randomized across class type by age.

4. Attrition Rate
There is a 2% chance of observing a sample such as this one or more extreme, given our null hypothesis is true. Being statistically significant, this means that class types may not be adequately randomized by attrition rate.

5. Class Size in Kindergarten
There is less than a 1% chance of observing a sample such as this one or more extreme, given our null hypothesis is true. Being statistically significant, this means that class types may not be adequately randomized by class size.

6. Percentile Score in Kindergarten
There is less than a 1% chance of observing a sample such as this one or more extreme, given our null hypothesis is true. Being statistically significant, this means that class type may not be adequately randomized by average SAT percentile.

b. Choose one of the characteristics whose P-value you decided was cause for concern in part (a) and explain briefly but clearly how the researchers dealt with this problem.

I would be most concerned about the three statistically significant items in which we can not be certain at an alpha level of 0.05 that we have adequately randomized our sample by class size, SAT score and attrition rate.

3. This question requires you to interpret data reported in the paper "Institutional Corruption and Election Fraud: Evidence from a Field Experiment in Afghanistan" (Callen and Long 2015).

   The authors study the impacts of a new monitoring technology on the manipulation of vote totals during the 2010 parliamentary elections in Afghanistan. Specifically, they test whether announcing the use of this technology at a given polling center to election officials reduces fraud. In the run up to elections, they deliver a letter explaining that the monitoring technology will be used at that location to a randomly selected set of polling center managers in 238 polling centers from an experimental sample of 471 polling centers.

   Table 9 examines the impacts of sending the letter on the votes received by the candidate with the strongest political connections. The data from four regressions for polling centers (PCs) are reproduced here, with standard errors in parentheses:

TABLE 9—SPATIAL TREATMENT EXTERNALITIES

	Votes for the most connected candidate			
	(1)	(2)	(3)	(4)
Letter treatment (= 1)	−4.080**	−4.183**	−4.290**	−4.159**
	(2.009)	(1.982)	(1.956)	(1.980)
Any PCs treated within 1 km (= 1)		−6.877*	−6.742*	
		(3.512)	(3.486)	
Total PCs within 1 km		−0.597	−0.499	−1.256
		(0.566)	(0.564)	(0.806)
Any PCs treated within 1–2 km (= 1)			−4.738	−4.681
			(4.244)	(4.240)
Total PCs within 1–2 km			0.103	0.223
			(0.378)	(0.392)
1 treated PC within 1 km (= 1)				−6.457*
				(3.613)
2 treated PCs within 1 km (= 1)				−5.831
				(3.882)
3 treated PCs within 1 km (= 1)				−3.007
				(4.858)
4 treated PCs within 1 km (= 1)				1.459
				(5.620)
5 treated PCs within 1 km (= 1)				−1.334
				(6.929)
Constant	28.064***	30.543***	32.378***	32.697***
	(6.017)	(6.043)	(7.004)	(6.987)
R^2	0.276	0.290	0.292	0.294
Trimming top 1 percent of votes for interacted candidate type	Yes	Yes	Yes	Yes
Number polling centers	439	439	439	439
Number candidate–polling substation observations	1,841	1,841	1,841	1,841
Mean dep. var. control + no treated PCs 0–2 km	42.939	42.939	42.939	42.939

a. Explain, briefly but precisely, the meaning of the coefficient estimate of -4.080 on *Letter treatment (=1)* in regression (1) in the context of this evaluation.

Those PCs who received letters announcing the monitoring technology received 4.080 fewer votes for the 1st most connected candidate.

b. Interpret the coefficient estimate of -6.877 on *Any PCs treated within 1km (=1)* [a dummy variable for whether any other centers within 1km of a given polling center received the letter treatment] in regression (2) in the context of this evaluation. What does it tell us in terms of spillovers/externalities?

PCs within 1km of any other PCs had 6.877 fewer votes for the 2nd most connected candidate. This really highlights the potential of spillover effects, i.e. the impact of multiple PCs within a small geographic area on the awareness of the new monitoring technology to be used.

c. What does your answer to part (b) imply about whether the coefficient estimate from part (a) is an underestimate or overestimate (or neither) of the total impact of the letter treatment?

Part (b) indicates that our coefficient estimate from part (a) was an underestimate of the treatment effect.

4. Find an article from a newspaper, magazine, or online news outlet in which the results of a randomized evaluation (preferably of a social program or policy, as opposed to a drug) are reported and discussed.

a. Please provide a link to the article or otherwise cite the source of the article.

Patient Navigation for Colonoscopy Completion: Results of an RCT (Link)

DeGroff, A., Schroy, I. C., Morrissey, K. G., Slotman, B., Rohan, E. A., Bethel, J., & ... Joseph, D. (2017). Research Article: Patient Navigation for Colonoscopy Completion: Results of an RCT. American Journal Of Preventive Medicine, 53363-372. doi:10.1016/j.amepre.2017.05.010

b. What are the treatment(s) and outcome(s) of interest for this evaluation?

Implemented largely via telephone, the exposure/independent/x/treatment variable of interest was the role and educational efforts of lay health navigators on the facilitation and completion rates of colonoscopies. The outcome/dependent/y variable of interest were colonoscopy completion rates (within 6 months).

c. Briefly describe the findings of the evaluation.

Overall, they found that colonoscopy completion was significantly higher for patients that received the treatment of navigation education and facilitation (61.1% compared to 53.2%, p=0.021. And the odds of

colonoscopy completion for navigated patients was 1.5 times greater than the control group of patients without navigated care.

d. What are three questions you could ask the evaluators to help you assess the internal validity of the evaluation?

1) With a well defined association between first-degree relatives having a family history of colon cancer on the risk of developing cancer, what is the rationale for not collecting and accounting for such a covariate?

2) All participants in this study, even as members of an underserved community, had/have access to primary care. How might this data differ from underserved peoples without access to primary care?

3) Though mentioned in your introduction: Screening rates were lower among those with lower incomes and education and people of Hispanic/Latino ethnicity" and despite acknowledging a study by the CDC using the BRFSS data indicating that "screening rates are greater among those who are married, employed, insured, have a usual source of care, and who have greater incomes and education and lower among Hispanics," this study did not report data in Table 2 on by education levels.

5. The following questions are included to help me get to know you a bit better and to inform choices of examples I use in class in the hopes that these examples will be relevant to your interests and experiences. These items are not scored.
   a. Do you have any experience with evaluating programs/policies or working on programs/policies that have been evaluated? If so, please briefly describe.

   My background is in Public Health. I have done a fair amount of program evaluation through assessing methodology, not necessarily in analyzing outcome data, as such is not always available. Though I respect the insight quantitative/statistical data can offer on the effectiveness of a program, and I do have introductory experience assessing public health interventions and data statistically, especially in LMICs and low-resource nations, my main interest is in data collection methods, quality of data collection, and research design. Having taken a year's worth biostatistics at the School of Public Health here at Brown, I

will say that the methodology and how outcome data is presented is just not jiving with me.

b. Do you have any sectoral interests or expertise (education, health, labor, etc.) or regional interests or expertise (international, domestic) that you'd like me to keep in mind as I choose examples to analyze in the course? If so, please share.

Again, public health. I got my master's here at Brown and am most interested in international programming. In terms of specific data interests, I have done some quantitative research on cancer and veteran status, qualitative research on disability and apparel design, and spatial data analysis on malaria prevalence accounting for infrastructural and environmental factors. Outside of academia, I have extensive experience with orthopedics, clinical research on health-outcomes for joint replacement and scoliosis surgeries.

Due on Canvas site by midnight Wednesday, October 18. You may work together in small groups, but each student must write up his or her answers separately and list the names of the other students with whom he or she worked.

INSTRUCTIONS FOR SUBMISSION: Please upload a single document to the Canvas site with your responses to the questions in this problem set. For Question 1, which includes work in STATA, please copy and paste the STATA commands you used and the STATA output each command generated for each part (a) through (f) of the question. Be sure to include answers to all of the questions in each part. Your responses to each of the 6 parts of Question 1 should include a line or two of STATA code/output as well as a few sentences that respond to the questions in that part.

The first part of this assignment asks you to analyze the same data that Dehejia and Wahba used in their 1999 paper. This paper is on the reading list, and it may be helpful to read it before beginning work on this part of the assignment.

<u>You have access to two STATA datasets</u>:

- `nsw_exper.dta` – Contains the data from the randomized evaluation of the National Supported Work (NSW) Demonstration, a labor training program. The dataset contains 445 observations, including 185 in the treatment group and 260 in the control group.

- `nsw_psid.dta` – Contains non-experimental data from the Population Survey of Income Dynamics (PSID). The dataset contains 2,490 observations, none of which were treated.

<u>The two datasets use the following variables</u>:

Treatment variable
- `nsw` indicator for participation in NSW (1 if participated, 0 otherwise)

Outcome variables
- `re78` real earnings for 1978
- `u78` indicator for employment status in 1978 (1 if employed, 0 otherwise)

Covariates

- `age` age in years
- `educ` years of education
- `black` indicator for race (1 if black, 0 otherwise)
- `hispanic` indicator for ethnicity (1 if Hispanic, 0 otherwise)
- `married` indicator for marital status (1 if married, 0 otherwise)
- `re75` real earnings for 1975
- `re74` real earnings for 1974
- `u75` indicator for employment status in 1975 (1 if employed, 0 otherwise)
- `u74` indicator for employment status in 1974 (1 if employed, 0 otherwise)

For this question, you will need to use the STATA command nnmatch (which stands for "nearest neighbor match"), which may not already be loaded onto your version of STATA. To ensure you can use this command, make sure your computer is connected to the Internet and type the following into the STATA command line:

```
net describe st0072
from(http://www.stata-journal.com/software/sj4-3)
```

Then type: `net install st0072`

To use the `nnmatch` command, use the following syntax:
nnmatch `depvar treatvar varlist, m(# matches) tc(att) robust(4) pop`

Where *depvar* is the dependent (or outcome) variable, *treatvar* is the indicator variable for the treatment, *varlist* is a list of all the covariates on which you want to match, and # *matches* is the number of matches you want to use. Including *tc(att)* estimates the average treatment effect on the treated. (The default is the
average treatment effect.) Including *robust(4)* and pop are used to specify how standard errors are calculated. Note that nnmatch can take a bit of time to finish running and present results.

For parts (c) – (f), you will need to merge the experimental and non-experimental datasets. To do this, run `merging.do` (the file is included in this assignment). You can run this file by typing the following into

the STATA command line:
`do merging`

This will create a merged dataset called `nsw_psid_withtreated.dta` that deletes the 260 observations from the experimental control group and replaces them with the 2,490 untreated observations from the non-experimental control group. Make sure that the three relevant files (`nsw_exper.dta`, `nsw_psid.dta`, and `merging.do`) are all in the STATA working directory before you do this merge. You can determine what the STATA working directory is on your computer by typing `pwd` in the command line.

Other useful commands for this assignment:

- `reg depvar varlist, robust` (linear regression of *depvar* on variables in *varlist*)
- `probit depvar varlist` (probit regression of *depvar* on variables in *varlist*)
- `predict ps` (when used after you run the probit regression, it will compute the probit estimate for each observation and store it in the variable *ps*)

******I worked with Yi Zeng on a good portion of this problem set.***

Question 1

a. Using only the experimental data, estimate the effect of NSW on 1978 earnings without controlling for any covariates. Explain why this estimate is an unbiased estimate of the causal effect of the program.

```
. regress re78 nsw

      Source |       SS       df       MS              Number of obs =     445
-------------+------------------------------           F(  1,   443) =    8.04
       Model |   348013455     1   348013455           Prob > F      =  0.0048
    Residual |  1.9178e+10   443  43290390.9           R-squared     =  0.0178
-------------+------------------------------           Adj R-squared =  0.0156
       Total |  1.9526e+10   444  43976704.1           Root MSE      =  6579.5

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         nsw |   1794.343   632.8536     2.84   0.005     550.5749    3038.111
       _cons |   4554.802    408.046    11.16   0.000     3752.856    5356.749
------------------------------------------------------------------------------
```

This estimate (beta coefficient) of $1,794.34 is an unbiased estimate of the causal effect of the NSW labor training program because this output indicates an estimate equal to the parameter being estimated. Essentially this estimate is unbiased because it estimates the causal effect of the population from which the sample was drawn. This estimate indicates that those in the treated group of the NSW labor training program earned an estimated $1,794.34 more than those in the control group.

b. Using only the experimental data, estimate the effect of NSW on 1978 earnings, this time controlling for age, education, race, ethnicity, marital status, and income and employment in 1974 and 1975 in a linear regression. Comment on this estimate compared to the estimate in part (a). Is this what you expected? Why or why not?

Source	SS	df	MS		
				Number of obs =	445
				F(10, 434) =	2.68
Model	1.1368e+09	10	113680282	Prob > F =	0.0034
Residual	1.8389e+10	434	42370630.9	R-squared =	0.0582
				Adj R-squared =	0.0365
Total	1.9526e+10	444	43976704.1	Root MSE =	6509.3

re78	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
nsw	1672.042	634.2895	2.64	0.009	425.381	2918.703
age	53.66766	45.30327	1.18	0.237	-35.37343	142.7088
educ	402.9471	177.4222	2.27	0.024	54.23361	751.6606
black	-2039.466	1163.826	-1.75	0.080	-4326.903	247.9704
hisp	424.6486	1560.83	0.27	0.786	-2643.077	3492.375
married	-146.6618	880.9651	-0.17	0.868	-1878.15	1584.827
re74	.1235727	.0869804	1.42	0.156	-.0473824	.2945278
re75	.0194585	.1489178	0.13	0.896	-.2732313	.3121483
u74	1380.999	1185.606	1.16	0.245	-949.2444	3711.242
u75	-1071.817	1022.922	-1.05	0.295	-3082.314	938.6797
_cons	221.4286	2632.962	0.08	0.933	-4953.513	5396.371

This new estimate, a beta coefficient of $1,672.04 is the estimate of the causal effect of the NSW labor training program controlling for the following covariates: age education black and hispanic ethnicity, marital status, income and employment in 1974. This estimate is $122.30 lower than the previous unbiased and uncontrolled

estimates and this was to be expected. I anticipated that without controlling for covariates initial estimates would be higher, and overestimate the causal effect.

For the following questions, use the merged dataset, nsw_psid_withtreated.dta (see instructions above).

c. Using the non-experimental controls, compute the difference in average earnings between NSW participants and nonparticipants. Compare this to the result in part (a). Why are these figures so different?

```
. net describe st0072, from(http://www.stata-journal.com/software/sj4-3)

------------------------------------------------------------------------
package st0072 from http://www.stata-journal.com/software/sj4-3
------------------------------------------------------------------------

TITLE
      SJ4-3 st0072.  Implementing matching estimators for average ...

DESCRIPTION/AUTHOR(S)
      Implementing matching estimators for average treatment
        effects in Stata
      by Jane Leber Herr, UC Berkeley
         David M. Drukker, StataCorp
         Guido W. Imbens, UC Berkeley
         Alberto Abadie, Harvard University
      Support:  herrjl@yahoo.com, ddrukker@stata.com
      After installation, type help nnmatch

INSTALLATION FILES                               (type net install st0072)
      st0072/nnmatch.ado
      st0072/nnmatch.hlp

ANCILLARY FILES                                  (type net get st0072)
      st0072/nnmatch.do
      st0072/artificial.dta
      st0072/ldw_exper.dta
------------------------------------------------------------------------

. net install st0072
checking st0072 consistency and verifying not already installed...
installing into c:\ado\plus\...
installation complete.
```

```
. pwd
C:\Users\ksarcone\Documents

. do merging

. /* Merge */
. clear

. set memory 250m
set memory ignored.
    Memory no longer needs to be set in
modern Statas; memory adjustments are
performed on the fly automatically.

. set matsize 800
```

```
. use nsw_exper

. keep if nsw==1
(260 observations deleted)

. save treated , replace
(note: file treated.dta not found)
file treated.dta saved

. use nsw_psid

. append using treated

. save nsw_psid_withtreated , replace
(note: file nsw_psid_withtreated.dta not
found)
file nsw_psid_withtreated.dta saved

. end of do-file
```

```
. save "C:\Users\ksarcone\Documents\nsw_psid_withtreated.dta", replace
file C:\Users\ksarcone\Documents\nsw_psid_withtreated.dta saved

. regress re78 nsw
```

Source	SS	df	MS
Model	3.9811e+10	1	3.9811e+10
Residual	6.1365e+11	2673	229573197
Total	6.5346e+11	2674	244375670

Number of obs	=	2675
F(1, 2673)	=	173.41
Prob > F	=	0.0000
R-squared	=	0.0609
Adj R-squared	=	0.0606
Root MSE	=	15152

re78	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
nsw	-15204.78	1154.614	-13.17	0.000	-17468.8	-12940.75
_cons	21553.92	303.6414	70.98	0.000	20958.53	22149.32

```
. sum re78

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        re78 |      2675    20502.38    15632.52          0   121173.6

. use "C:\Users\ksarcone\Documents\nsw_exper.dta", clear

. sum re78

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
        re78 |       445    5300.765    6631.493          0    60307.9
```

The estimated difference in average earnings between NDW participants and nonparticipants is $15,204.78 This figure is very different from part (a) because it is only considering non-experimental controls in which the average earnings is much higher ($20,502.38) than those only in the experimental group ($5,300.765) and the sample sizes vary between the two. In part (a) we have a control group of 260 and with part (c) we have a control group sample of 2,490

d. Use the same linear regression as in part (b) to control for all the covariates, this time with the non-experimental data as the comparison group. Can you replicate the experimental results from part (a) by controlling for all these observed differences between participants and nonparticipants? If not, why might that be?

```
. regress re78 nsw age educ black hisp married re74 re75 u74 u75

      Source |       SS       df       MS              Number of obs =    2675
-------------+------------------------------           F( 10,  2664) =  378.79
       Model |  3.8364e+11    10  3.8364e+10           Prob > F      =  0.0000
    Residual |  2.6982e+11  2664   101282294           R-squared     =  0.5871
-------------+------------------------------           Adj R-squared =  0.5855
       Total |  6.5346e+11  2674   244375670           Root MSE      =   10064

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         nsw |   115.3825   1006.884     0.11   0.909    -1858.972    2089.737
         age |  -89.76542   21.93963    -4.09   0.000    -132.7858   -46.74499
        educ |    514.124   76.44391     6.73   0.000     364.2285    664.0194
       black |   -454.216   496.8819    -0.91   0.361    -1428.529    520.0974
        hisp |   2197.373   1091.634     2.01   0.044     56.83747    4337.908
     married |   1204.785   585.4794     2.06   0.040     56.74446    2352.825
        re74 |     .31262   .0316311     9.88   0.000      .250596    .3746441
        re75 |   .5436544   .0309038    17.59   0.000     .4830566    .6042522
         u74 |   2389.531   1024.439     2.33   0.020     380.7544    4398.307
         u75 |  -1461.965   947.1953    -1.54   0.123    -3319.278    395.3475
       _cons |   953.6012   1370.579     0.70   0.487    -1733.904    3641.107
------------------------------------------------------------------------------
```

No, we can not reproduce the experimental results from part (a) because in part a we were dealing with 185 treated and 260 experimental control observations and did not control from any covariates. In this question we are looking at the same 185 treated units against 2490 non-experimental controls, AND were control for listed covariates.

e. Use the nnmatch command to estimate the average effect of the NSW program on the treated, matching on all the covariates and using 1, 4, 10, and 20 matches. (Note that this will mean four separate estimates, one for each choice of # matches.) How do these estimates compare to the experimental estimate in part (a)? Explain why each of the four estimates is so different from the other three.

```
. nnmatch re78 nsw age educ black hisp married re74 re75 u74 u75, m(1) tc (att) robust(4) pop

Matching estimator: Population Average Treatment Effect for the Treated

Weighting matrix: inverse variance          Number of obs          =      2675
                                            Number of matches  (m) =         1
                                            Number of matches,
                                              robust std. err. (h) =         4

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        PATT |   2073.478   1128.281     1.84   0.066    -137.9118    4284.868
------------------------------------------------------------------------------
Matching variables:  age educ black hisp married re74 re75 u74 u75
```

This estimate is $279.14 more than part (a).

```
. nnmatch re78 nsw age educ black hisp married re74 re75 u74 u75, m(4) tc (att) robust(4) pop

Matching estimator: Population Average Treatment Effect for the Treated

Weighting matrix: inverse variance          Number of obs          =      2675
                                            Number of matches  (m) =         4
                                            Number of matches,
                                              robust std. err. (h) =         4

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        PATT |   1618.731   911.6289     1.78   0.076    -168.0285    3405.491
------------------------------------------------------------------------------
Matching variables:  age educ black hisp married re74 re75 u74 u75
```

This estimate is $175.61 less than part (a).

```
. nnmatch re78 nsw age educ black hisp married re74 re75 u74 u75, m(10) tc (att) robust(4) pop

Matching estimator: Population Average Treatment Effect for the Treated

Weighting matrix: inverse variance          Number of obs          =     2675
                                            Number of matches  (m) =       10
                                            Number of matches,
                                              robust std. err. (h) =        4

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        PATT |   1259.345   865.3125     1.46   0.146    -436.6362    2955.326
------------------------------------------------------------------------------
Matching variables:  age educ black hisp married re74 re75 u74 u75
```

This estimate is $535.00 less than part (a).

```
. nnmatch re78 nsw age educ black hisp married re74 re75 u74 u75, m(20) tc (att) robust(4) pop

Matching estimator: Population Average Treatment Effect for the Treated

Weighting matrix: inverse variance          Number of obs          =     2675
                                            Number of matches  (m) =       20
                                            Number of matches,
                                              robust std. err. (h) =        4

------------------------------------------------------------------------------
        re78 |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        PATT |   351.1535   832.5011     0.42   0.673    -1280.519    1982.826
------------------------------------------------------------------------------
Matching variables:  age educ black hisp married re74 re75 u74 u75
```

This estimate is $1,443.19 less than part (a).

"nnmatch estimates the average treatment effect on depvar by comparing outcomes between treated and control observations, using nearest neighbor matching across [selected] variables…[1]"

Since in many cases perfect matches are not available, nearest neighbor allows us to estimate average treatment effect on the outcome variable while looking at matches of nearest neighbors for treated and control observations. Reasonably we find that

the more number of matches the smaller the estimated effect, which is why we see a notable difference between one match $2,073.48 and 20 matches $351.15

f. Estimate the propensity score using a probit regression. Use the nnmatch command to estimate the average effect of the NSW program on the treated by matching only on the propensity score and using just 1 match. Compare this to the result in part (a). How does this estimate perform (with respect to the experimental estimate) compared to those in parts (c) – (e)?

```
. probit re78 nsw age educ black hisp married re74 re75 u74 u75

Iteration 0:   log likelihood = -1001.2732
Iteration 1:   log likelihood = -700.48076
Iteration 2:   log likelihood = -690.81336
Iteration 3:   log likelihood = -690.76597
Iteration 4:   log likelihood = -690.76597

Probit regression                                 Number of obs   =       2675
                                                  LR chi2(10)     =     621.01
                                                  Prob > chi2     =     0.0000
Log likelihood = -690.76597                       Pseudo R2       =     0.3101
```

re78	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
nsw	.5414877	.1592108	3.40	0.001	.2294404	.8535351
age	-.0181056	.0040529	-4.47	0.000	-.0260492	-.0101621
educ	-.00385	.0142568	-0.27	0.787	-.0317927	.0240927
black	.1489637	.0964222	1.54	0.122	-.0400203	.3379477
hisp	.6033141	.2460885	2.45	0.014	.1209895	1.085639
married	.0951244	.1066439	0.89	0.372	-.1138938	.3041426
re74	-1.18e-06	6.07e-06	-0.19	0.846	-.0000131	.0000107
re75	.000034	6.36e-06	5.34	0.000	.0000215	.0000465
u74	-.7390463	.1546321	-4.78	0.000	-1.04212	-.4359729
u75	-.7552845	.1418931	-5.32	0.000	-1.03339	-.4771791
_cons	1.53735	.2594371	5.93	0.000	1.028863	2.045837

A propensity score is essentially the probability of being in the treatment group given the observed values of the covariates. The propensity score is 54%

The two step process for propensity score matching involves estimating the propensity score using the probit regression and then estimating the average treatment effect using the nnmatch command in STATA.

```
. nnmatch re78 nsw ps, m(1) tc (att) robust(4) pop

Matching estimator: Population Average Treatment Effect for the Treated

                                       Number of obs           =      2675
                                       Number of matches   (m) =         1
                                       Number of matches,
                                         robust std. err. (h) =         4
```

re78	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
PATT	-9047.703	2034.701	-4.45	0.000	-13035.64	-5059.763

```
Matching variables:  ps
```

When only matching on the propensity score, the average treatment effect of the NSW labor training program is -$9,047.70 translating to a difference of $10,842.04 in the estimated effect from part (a).

This nnmatch estimate provides the average treatment effect on `re78` / real earnings for 1978 by comparing outcomes between treated and control observations using only 1 nearest neighbor match across propensity score.

This estimate is closest to the previous experimental estimate of part (c) however, this estimate is drastically different than those in which we controlled for all covariates in part (d) and using nearest neighbor matches, part (f).

Question 2

Workers' compensation programs provide payments for medical care and cash benefits for work-related injuries. You want to estimate the effect of high benefits on the duration of claims. The question is whether high benefits induce workers to stay out of work longer to complete medical recovery or to have more leisure. Kentucky recently raised the benefit amount for high earnings individuals by almost 50 percent; neighboring Tennessee did not raise the benefit amount. Data on the average duration (in weeks) of temporary disabilities for high earnings individuals before and after the change in Kentucky are reported here:

Duration of Temporary Disabilities for High Earnings Individuals

	Before Increase	After Increase
Kentucky [Intervention]	10.5	14.0
Tennessee [Comparison]	8.7	10.2

a. Using a difference-in-differences (DD) estimator, compute an estimate of the effect of the higher benefit amount on the average duration of temporary disabilities.

Controlling for observed characteristics doesn't address any influences that may have affected the treated and untreated individuals differently in unobservable or measured ways. Thus we can measure the difference-in-differences (DD). Through pre-post / simple differences we can compare the changes in outcome between a treated and control group.

Subtracting out "systematic" differences using pre-intervention data:

	Before	After
Intervention [KY]	10.5 [T1]	14.0 [T2]
Comparison [TN]	8.7 [C1]	10.2 [C2]
Difference	1.8 [T1-C1]	3.8 [T2-C2]

[T2 - C2] - [T1 - C1] = 3.8 - 1.8 = 2.0

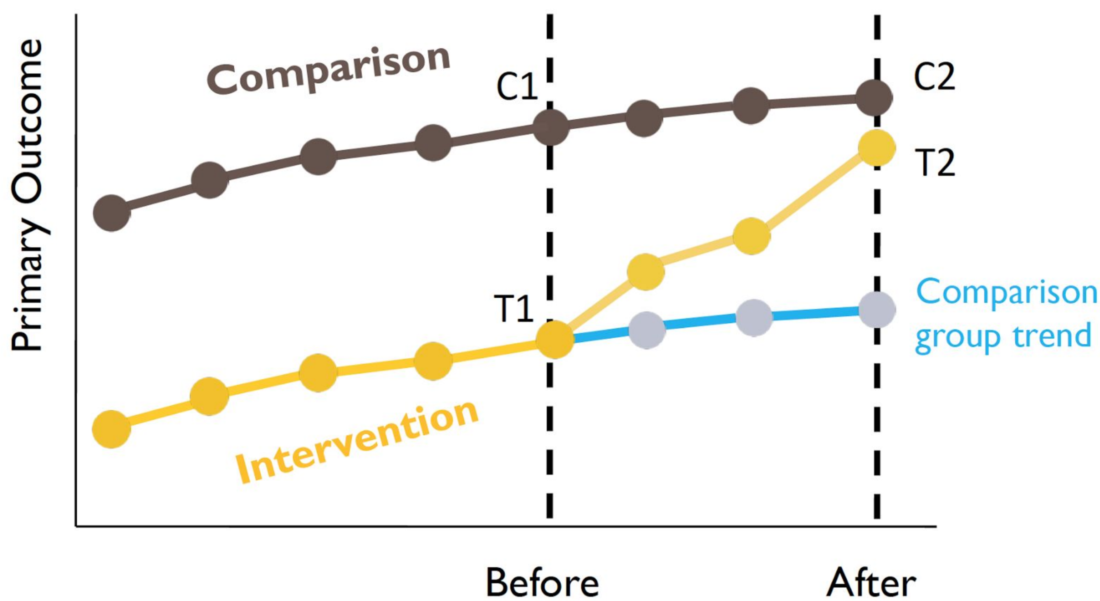
Subtracting out "trend" from comparison group:

	Before	After	Difference
Intervention [KY]	10.5 [T1]	14.0 [T2]	3.5 [T2-T1]
Comparison [TN]	8.7 [C1]	10.2 [C2]	1.5 [C2-C1]

[T2 - T1] - [C2 - C1] = 3.5 - 1.5 = 2.0

b. What assumption is necessary for this estimate to be unbiased? How might you test the plausibility of that assumption?

In order for this estimates to be unbiased we must assume that the "intervention and comparison groups would have the same outcome in the absence of the program and that the outcome variable would have remained constant in the absence of the program[2]". See image below from Lecture 9:



Because we can't <u>know</u> how the world would be different without the program the best we could possibly do to assess validity would involve checking pre-intervention trends, having a placebo intervention group (a fake intervention group not affected

by the program) and a placebo outcome (something that can't logically be affected by the program). You'll be looking to see if both intervention and comparison groups have moved in parallel before the program started, a trend that may be likely or destined to continue in absence of the program. And you'll want to calculate that the difference-in-differences for the placebo intervention group and the placebo outcome is very close to zero.

Furthermore, with more data or state options you may want to make sure that states that closely resemble the intervention group as best as possible and control for state fixed effects. It may be that TN is not the best comparison group for KY if given more data statistics around age, disability, employment, access to medical care etc. were not as similar to KY as OK.

Below are the same data for low earnings individuals, who were not affected by the increase in benefits for high earnings individuals in Kentucky.

Duration of Temporary Disabilities for Low Earnings Individuals

	Before Increase	After Increase
Kentucky	9.6	11.8
Tennessee	7.1	7.5

c. Discuss how these data on low earnings individuals affect your confidence in the DD estimate you computed in part (a). How might you use these additional data to refine the estimate?

Subtracting out "systematic" differences using pre-intervention data:

	Before	After
Intervention [KY]	9.6 [T1]	11.8 [T2]
Comparison [TN]	7.1 [C1]	7.5 [C2]
Difference	2.5 [T1-C1]	4.3 [T2-C2]

[T2 - C2] - [T1 - C1] = 4.3 - 2.5 = 1.8

Subtracting out "trend" from comparison group:

	Before	After	Difference
Intervention [KY]	9.6 [T1]	11.8 [T2]	2.2 [T2-T1]
Comparison [TN]	7.1 [C1]	7.5 [C2]	0.4 [C2-C1]

[T2 - T1] - [C2 - C1] = 2.2 - 0.4 = 1.8

The fact that there is only a 0.2 difference (in days) between the DD estimate from part (a) and part (c) really negatively impacts my confidence that there is any meaningful effect happening in high earning individuals receiving additional support on temporary disability/workers' compensation, post benefit adjustments.

You could use this additional data of low-income earners as a placebo intervention since they did not receive the same recent raise in benefits of almost 50% as the high earning individuals.

Question 3

A microfinance institution (MFI) operating in rural India is interested in measuring the impact of an individual receiving one of its loans on the future income of the individual. The MFI offers its loans to all qualified borrowers in the region, though not all potential borrowers do in fact take a loan. The MFI observes which individuals in the region takes a loan and then uses gender, age, pre-loan income, occupation, and family size to compute a propensity score for a random sample of individuals in the region, some who have taken a loan and some who have not. A subset of the data collected, including income two years after being offered (and potentially receiving) a loan, is show here:

Individual ID	Received a Loan?	Income (2yrs after offer)	Propensity Score
1	No	$2,000	0.95
2	No	$3,000	0.70
3	Yes	$2,000	0.90
4	Yes	$5,000	0.85

5	No	$3,500	0.60

a. What is the meaning of these propensity scores? Be precise.

A propensity score is essentially the probability of being in the treatment group given the observed values of chosen covariates. Given this data, the propensity score for each individual is the probability of accepting the initial loan given the covariates of gender, age, pre-loan income, occupation and family size. For example for Individual ID #1, they were 95% likely to accept the loan given their age, gender, pre-loan income, occupation and family size.

$p(X) = Pr(D = 1|X)$

b. Using only these data, use propensity score matching (inexact) with replacement to compute an estimate of the average treatment effect on the treated (ATET), where the treatment is receiving a loan and the outcome of interest is income after two years.

When perfect matches are not available or feasible we use nearest neighbor for inexact matching. In this case we do not have the option for exact matching. Inexact matching using propensity scores:

Treated ID & Propensity Score	Nearest Neighbor Control ID & Propensity Score
[3] = 0.90 $2,000	[1] = 0.95 $2,000
[4] = 0.85 $5,000	[2] = 0.70 $3,000

***Control [5] is dropped**

Y = outcome/dependent variable = income after 2 years
X = exposure/treatment = receiving a loan

$$= \sum \ [(E[\ Y \mid X , D = 1] - E[Y \mid X , D = 0]) \times Pr (X \mid D = 1)]$$
$$(\$3{,}500 - \$2{,}500) \times 0.50$$
$$ATET = \$500$$

c. Aside from the small sample size, explain one other concern you have with regard to the validity of this estimate. Be specific.

Firstly, our analysis is only as good as our data and so it's hard to consider characteristics that composed the pre-program make up that were not measured or observed. This is one of the major limitations of propensity scores. Additionally, we're using post-intervention characteristics for matching, which is problematic. Also group overlap must be substantial to ensure adequate or quality matching. Furthermore propensity score analysis (PSA) or propensity score matching (PSM) does not take into account clustering and sometimes exacerbates imbalance[3]. Lastly, because we are doing inexact matching with replacement, and controls can be used more than once we have higher variance.

Endnotes:

[1] Abadie, A., D. Drukker, J. L. Herr, and G. W. Imbens. (2004). Implementing matching estimators for average treatment effects in Stata. Stata Journal 4(3): 290-311.

[2] Neggers, Yousef. (2017). Matching Estimators 2 Diff in Diff 1. [PowerPoint slides]. Retreived from Canvas 10.12.2017

[3] Gary King and Richard Nielsen. Working Paper. "Why Propensity Scores Should Not Be Used for Matching". Copy at http://j.mp/2ovYGsW

Question 1

a. (*1 point*) Use the `reg` command to estimate the effect of being assigned to the training on earnings for men and for women (separately). This is the Intent-to-Treat (ITT) effect.

```
. reg earnings assigned if female == 1, robust

Linear regression                               Number of obs     =        400
                                                F(1, 398)         =     167.16
                                                Prob > F          =     0.0000
                                                R-squared         =     0.2578
                                                Root MSE          =     1767.7
```

earnings	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
assigned	2400	185.6289	12.93	0.000	2035.064	2764.936
_cons	6000	152.1239	39.44	0.000	5700.933	6299.067

```
. reg earnings assigned if female == 0, robust

Linear regression                               Number of obs     =        600
                                                F(1, 598)         =       1.80
                                                Prob > F          =     0.1808
                                                R-squared         =     0.0030
                                                Root MSE          =     1828.1
```

earnings	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
assigned	200	149.2668	1.34	0.181	-93.15079	493.1508
_cons	9800	116.3957	84.20	0.000	9571.406	10028.59

This question looking at the Intent-to-Treat (ITT) effect tells us the causal effect of the "offer" of treatment, knowing that some observations that will be offered treatment may decline participation.

$$ITT = E[Y|Z = 1] - E[Y|Z = 0]$$

This is important to consider when scaling up programs. We see that this is true later in which not all assignees enroll in the treatment.

The ITT for Males is $200 & $2,400 for Females. Meaning that females and males assigned to the treatment group respectively earned an average of $2,400 and $200 more than females and males that were placed in the control group.

b. (*1 point*) Use the `ivregress` command to estimate the effect of enrolling in training on earnings for men and for women (separately). This is the Local Average Treatment Effect (LATE).

```
. ivregress 2sls earnings (enrolled = assigned) if female == 1, robust

Instrumental variables (2SLS) regression          Number of obs   =        400
                                                  Wald chi2(1)    =     142.78
                                                  Prob > chi2     =     0.0000
                                                  R-squared       =          .
                                                  Root MSE        =     2114.6

------------------------------------------------------------------------------
             |               Robust
    earnings |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    enrolled |       3000   251.0617    11.95   0.000     2507.928    3492.072
       _cons |       6000   151.7431    39.54   0.000     5702.589    6297.411
------------------------------------------------------------------------------
Instrumented:  enrolled
Instruments:   assigned
```

```
. ivregress 2sls earnings (enrolled = assigned) if female == 0, robust

Instrumental variables (2SLS) regression          Number of obs   =        600
                                                  Wald chi2(1)    =       1.82
                                                  Prob > chi2     =     0.1772
                                                  R-squared       =     0.0138
                                                  Root MSE        =     1815.1

------------------------------------------------------------------------------
             |               Robust
    earnings |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    enrolled |        400   296.4104     1.35   0.177    -180.9538    980.9538
       _cons |       9800   116.2015    84.34   0.000     9572.249    10027.75
------------------------------------------------------------------------------
Instrumented:  enrolled
Instruments:   assigned
```

This question, using instrumental variables, estimates the Local Average Treatment Effect (LATE) for compliers.

$$\boxed{LATE_{compliers} = E[Y_1 - Y_0 | D_1 > D_0] =}$$

$$\boldsymbol{LATE = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}}$$

Essentially the average effect of offering the training program (treatment) for males was $400 & $3,000 for females. Meaning that males that complied/took the treatment and enrolled in the training program earned $400 more than those who did not take the treatment/did not enroll. Likewise, females that complied/took the treatment and enrolled in the training program earned $3,000 more than those who did not take the treatment/did not enroll.

c. (*1 point*) Explain why in this case the estimate from (b) can also be considered the ATET.

The estimate/LATE from part (b) can also be considered the ATET because no one assigned to the control group had the option to enroll in the treatment/training program, meaning there were essentially no always-takers. This is an example of one-sided compliance. And we can guarantee that no one in the control group had the training because the administrators of the experiment are the only providers.

$$\boxed{LATE_{compliers} = E[Y_1 - Y_0|D_1 > D_0] = E[Y_1 - Y_0|D = 1] = ATET}$$

d. (*1 point*) Complete the following tables separately for men and women in this dataset:

```
. sum if female == 1 & assigned == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        300    477.2867    285.7825          1        996
      female |        300           1           0          1          1
    earnings |        300        8400    1840.991       5298      11597
    assigned |        300           1           0          1          1
    enrolled |        300          .8    .4006683          0          1
```

```
. sum if female == 1 & assigned == 0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        100      521.94    291.8234          2       1000
      female |        100           1           0          1          1
    earnings |        100        6000    1525.075       4243       9159
    assigned |        100           0           0          0          0
    enrolled |        100           0           0          0          0
```

```
. sum if female == 0 & assigned == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        300    516.7233    292.0441          6        998
      female |        300           0           0          0          0
    earnings |        300       10000    1618.574       6047      12462
    assigned |        300           1           0          1          1
    enrolled |        300          .5    .5008354          0          1
```

```
. sum if female == 0 & assigned == 0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        300    500.3433    287.3839          4        997
      female |        300           0           0          0          0
    earnings |        300        9800    2016.032       6252      12866
    assigned |        300           0           0          0          0
    enrolled |        300           0           0          0          0
```

```
sum if female == 0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        600    508.5333    289.5975          4        998
      female |        600           0           0          0          0
    earnings |        600        9900     1829.35       6047      12866
    assigned |        600          .5    .5004172          0          1
    enrolled |        600         .25     .433374          0          1

sum if female == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        400      488.45    287.5891          1       1000
      female |        400           1           0          1          1
    earnings |        400        7800    2049.295       4243      11597
    assigned |        400         .75     .433555          0          1
    enrolled |        400          .6    .4905115          0          1
```

MEN	Assignees	Non-Assignees	Total
# Observations	300	300	600
# Enrolled in Training	150	0	150
Averaged Earnings	$10,000	$9,800	$9,900

FEMALE	Assignees	Non-Assignees	Total
# Observations	300	100	400
# Enrolled in Training	240	0	240
Averaged Earnings	$8,400	$6,000	$7,800

e. (*1 point*) Using the numbers from the tables you completed in (d) and basic arithmetic, show how you can replicate the ITT and ATET estimates you found using STATA in (a) and (b), again separately for men and women.

$$ITT = E[Y|Z=1] - E[Y|Z=0]$$

Men = $10,000 - $9,800 = $200
Women = $8,400 - $6,000 = $2,400

$$ATET = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}$$

$$Male\ ATET = \frac{\$10,000 - \$9,800}{\left(\frac{150}{300}\right) - \left(\frac{0}{300}\right)} = \frac{200}{0.50} = \$400$$

$$Female\ ATET = \frac{\$8,400 - \$6,000}{\left(\frac{240}{300}\right) - \left(\frac{0}{100}\right)} = \frac{2,400}{0.80} = \$3,000$$

f. (*1 point*) Discuss the conditions under which the ATET estimates from (b) and (e) can be considered unbiased estimates of the effect of the training program on earnings. For each, assess how likely you think the condition is to hold in this case.

For instrumental variables to yield unbias effect estimates there are four assumptions that must hold true. They include independence, exclusion, relevance and monotonicity. In this instance I believe all four assumptions hold true.

- **For independence I believe the assumption holds true because of the randomization methods used in this study for assignment (with females having a probability of assignment = 0.75, and males 0.5), translating into the instrument Z (being assigned or not assigned to a treatment or control group Z=1 or Z=0) is independent of potential outcomes and potential treatments.**
- **For exclusion I believe the assumption holds true because Z (assignment) only effects the outcome Y/earnings through its effect on D/enrollment/treatment in training program meaning there's a unique channel for causal effect on the outcome of interest. Essentially to be enrolled (D) you have to be assigned (Z), and assignment impacts enrollment probability.**
- **For relevance/first stage I believe the assumption holds true because the probability of enrolling does not equal the probability of not enrolling. Implying that the instrument (assignment) induces/causes this variation on D (enrollment). This is evident from looking at the differential take-up between treatment and control groups.**

$$\boxed{P(D_1 = 1) \neq P(D_0 = 1)}$$

- **For monotonicity I believe the assumption holds true because D1/enrolled is larger than being unenrolled/D0 and earnings are greater among enrolled. Also there are no defiers.**

g. (*1 point*) Aggregate the ATET estimates for men and women to estimate the overall average effect of the treatment on the treated in this study. (Hint: Notice that the proportions of men and women among the treated are *not* the same as the proportions of men and women in the whole sample.)

```
. ivregress 2sls earnings (enrolled = assigned), robust

Instrumental variables (2SLS) regression          Number of obs     =      1,000
                                                  Wald chi2(1)      =       5.54
                                                  Prob > chi2       =     0.0186
                                                  R-squared         =          .
                                                  Root MSE          =     2191.2

------------------------------------------------------------------------------
             |               Robust
    earnings |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    enrolled |   538.4615   228.7924     2.35   0.019     90.03672    986.8864
       _cons |       8850   125.7108    70.40   0.000     8603.611    9096.389
------------------------------------------------------------------------------
Instrumented:  enrolled
Instruments:   assigned
```

```
. sum if assigned == 1

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        600     497.005    289.3627          1        998
      female |        600          .5    .5004172          0          1
    earnings |        600        9200    1908.027       5298      12462
    assigned |        600           1           0          1          1
    enrolled |        600         .65    .4773676          0          1

. sum if assigned == 0

    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
          id |        400    505.7425     288.285          2       1000
      female |        400         .25     .433555          0          1
    earnings |        400        8850    2517.364       4243      12866
    assigned |        400           0           0          0          0
    enrolled |        400           0           0          0          0
```

Men & Women	D=1 / Enrolled	D=0 / Not Enrolled	Total	Earnings
Z=1 / Assigned	390	210	600	$9,200
Z=0 / Not-Assigned	0	400	400	$8,850

$$ATET = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]}$$

$$ATET = \frac{\$9,200 - \$8,850}{\left(\frac{390}{600}\right) - \left(\frac{0}{400}\right)} = \frac{350}{0.65} = \$538.46$$

Using both STATA and the LATE/ATET formula I found the aggregated average treatment effect on the treated in this study (both males and females) to be $538.46

Question 2

One of the most remarkable features of electoral politics in the United States is the high degree of electoral success of incumbent candidates. For example, for the last five decades, conditional on running for re-election, incumbent candidates have won elections to the House of Representatives about 90 percent of the time. This phenomenon has prompted much empirical research aimed to estimate the "incumbency advantage," that is, the *causal* effect of incumbency on the vote share obtained by a candidate.

Suppose that you have data on every two-candidate race (with both a Democrat and Republican running) to the House of Representatives for some period of time. You want estimate the "incumbency advantage" for Democratic candidates. Discuss the assumptions behind the following estimators, the validity of those assumptions, and the biases that will affect the estimators if the assumptions fail to hold (*1 point each*):

a. The "Vote Share Difference." The difference in average vote shares between Democratic incumbents and Democratic non-incumbents, the following omitted variables may be cause for bias concern:

   Here we're assuming the vote share difference portrays incumbent advantage. It's assuming that incumbents and challengers are identical in almost every other way than merely being an incumbent and challenger. But it doesn't account for attributes related to challenging candidate, districts, level of election and so forth.

   If you're assessing incumbency advantage through looking at the estimator of "vote share difference" between democratic incumbents and democratic non-incumbents, there is concern for Omitted Variable Bias, particularly with regards to:

   a. **The "quality" of candidates:**
      i. **High quality candidates are thought to more often run in open-seat elections, strategically saving their resources for more likely wins.**
      ii. **Relatedly, low-quality candidates are thought to run against incumbents more often than high-quality.**

iii. **This is thought of as the "scare-off" effect**

b. **Prior elected position of challenging candidates:**
   i. **Candidates that have held previous elected positions get higher votes than those who have not, and influences "quality"**

c. **Level of Election**
   i. **Different electoral cycles influence incumbency differently**

d. **Year Fixed Effects**

e. **Other Challenger Traits**
   i. **Ability to wait and afford for future elections and that influencing decision to run as a challenger vs. in an open seat election.**
   ii. **Gender, Race, Age, Symmetry of Face, Attractiveness, Business Experience, Leadership experience in the community**
      1. **These may also be influenced by time-fixed effects in which gender and race of challengers may matter more in incumbent races.**
   iii. **Income and Resource of challenging candidate**
      1. **Incumbents are considered to have added opportunities and resources to win votes, this is also likely more of an issue in more recent years as the cost of running a campaign has grown substantially, and it's believed to have influenced the quality of candidates running in incumbent races.**

f. **Seniority/Years in office**
   i. **The longer you're in office the more advantage you have as an incumbent**

g. **Voter turnout**

h. **Term Limits**
   i. **Impacts the probability of higher quality challengers**

i. **Incumbent Vulnerability**
   i. **Including political climate**

j. **State Fixed Effects**

k. **Previous Elections**
   i. **If previous elections were close seems impact incumbency differently than wider margin wins**

l. **Year/Temporal Factors**
   i. **Incumbency advantage has changed over time, growing from the 1950s- said to be related to increase in cost of running and the decrease in quality of challenging candidates**

b. The "Sophomore Surge." Among those Democratic candidates who were elected to office and who run for a second term immediately after the first term, the difference in average vote shares between the two elections. (That is, the average vote share gains for freshmen winners who run again in the following election.)

Using sophomore surge to estimate incumbency advantage assumes that no other items influence incumbency votes other than time in office could influence incumbency votes.

If you're assessing incumbency advantage through looking at the estimator of "Sophomore Surge" there is concern for Omitted Variable Bias, particularly with regards to:

a. **The quality of challenging candidate.**

    i. **High quality candidates are thought to more often run in open-seat elections, strategically saving their resources for more likely wins.**
    ii. **Relatedly, low-quality candidates are thought to run against incumbents more often than high-quality.**
    iii. **This is thought of as the "scare-off" effect**
  b. **Prior elected position of challenging candidates:**
    i. **Candidates that have held previous elected positions get higher votes than those who have not, and influences "quality"**
  c. **District:**
    i. **If the district is known to be a safely "Democratic" vs. contested**
      1. **Incumbents more often run in safe districts**
  d. **Year Fixed Effects**
  e. **Other Challenger Traits**
    i. **Ability to wait and afford for future elections and that influencing decision to run as a challenger vs. in an open seat election.**
    ii. **Gender, Race, Age, Symmetry of Face, Attractiveness, Business Experience, Leadership experience in the community**
    iii. **Income and Resource of challenging candidate**
      1. **Incumbents are considered to have added opportunities and resources to win votes, this is also likely more of an issue in more recent years as the cost of running a campaign has grown substantially, and it's believed to have influenced the quality of candidates running in incumbent races.**
  f. **Term Limits**
    i. **Impacts the probability of higher quality challengers**
  g. **Incumbent Vulnerability**
    i. **Including political climate**
  h. **State Fixed Effects**
  i. **Previous Elections**
    i. **If previous elections were close seems impact incumbency differently than wider margin wins**
  j. **Year/Temporal Factors**
    i. **Incumbency advantage has changed over time, growing from the 1950s- said to be related to increase in cost of running and the decrease in quality of challenging candidates**

c. The "Retirement Slump." In districts where the Democratic candidates retire, the difference in average vote shares obtained by the retiring Democratic incumbents in the previous election and the incoming Democratic candidates (who clearly are not incumbents) in the current election.

Here a dangerous elimination of data includes the elimination of democratic retirees that were replaced by republican candidates.

If you're assessing incumbency advantage through looking at the estimator of "Retirement Slump" there is concern for Omitted Variable Bias, particularly with regards to:

  a. **The quality of challenging candidate.**

i. High quality candidates are thought to more often run in open-seat elections, strategically saving their resources for more likely wins.
ii. Relatedly, low-quality candidates are thought to run against incumbents more often than high-quality.
iii. This is thought of as the "scare-off" effect

b. Prior elected position of challenging candidates:
- **i. Candidates that have held previous elected positions get higher votes than those who have not, and influences "quality"**

c. District:
- **i. If the district is known to be a safely "Democratic" vs. contested**
  - **1. Incumbents more often run in safe districts**

d. National vote/partisan swing

e. Level of Election
- **i. Different electoral cycles influence incumbency differently**

f. Year Fixed Effects

g. Other Challenger Traits
- **i. Ability to wait and afford for future elections and that influencing decision to run as a challenger vs. in an open seat election.**
- **ii. Gender, Race, Age, Symmetry of Face, Attractiveness, Business Experience, Leadership experience in the community**
- **iii. Income and Resource of challenging candidate**
  - **1. Incumbents are considered to have added opportunities and resources to win votes, this is also likely more of an issue in more recent years as the cost of running a campaign has grown substantially, and it's believed to have influenced the quality of candidates running in incumbent races.**

h. Lagged Vote Share

i. Lagged Party Control

j. Seniority/Years in office
- **i. The longer you're in office the more advantage you have as an incumbent**

k. Term Limits
- **i. Impacts the probability of higher quality challengers**

l. Incumbent Vulnerability
- **i. Including political climate**

m. State Fixed Effects

n. Party of Sitting Presidency

o. Previous Elections
- **i. If previous elections were close seems impact incumbency differently than wider margin wins**

p. Year/Temporal Factors
- **i. Incumbency advantage has changed over time, growing from the 1950s- said to be related to increase in cost of running and the decrease in quality of challenging candidates**

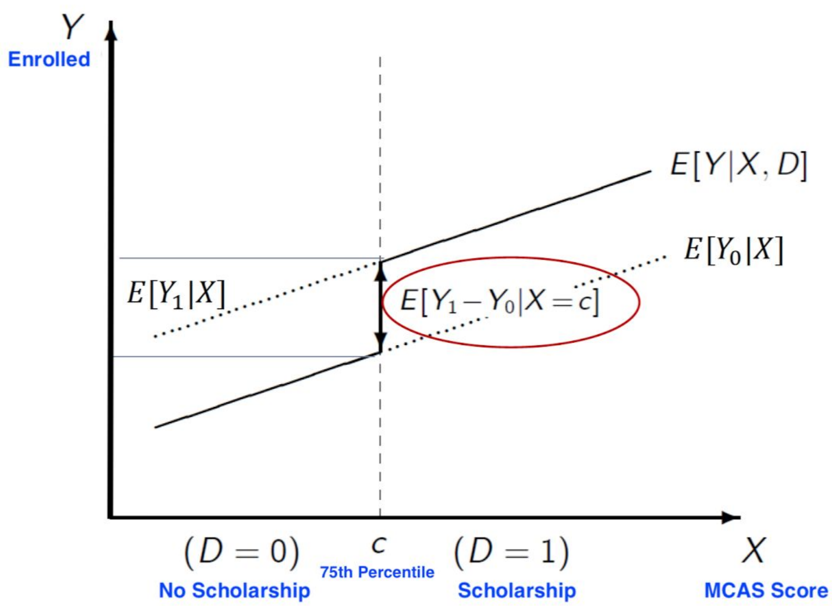
Question 1

In 2004, Massachusetts unveiled the Adams Scholarship program, an initiative designed to keep talented graduating high school seniors in-state for college. Under the program, students who scored in the top one-quarter of the state on the MCAS, a statewide standardized test, would get four tuition-free years at any in-state public college.

a. (*1 point)* Describe how evaluators might use a regression discontinuity design (RDD) to measure the effect scholarship eligibility has on the likelihood a student enrolls in an in-state public college. Be specific, and include how you would interpret the effect estimates this design would generate.

In this scenario the Adams Scholarship is the treatment (D), the "running variable," X is the score on the MCAS, and the threshold (c) is the top 25%/75th percentile. And the outcome (Y) is the likely to enroll in an in-state public college. Usually X is correlated with outcome Y, but in this case that may not necessarily be true, especially if it's more likely that smarter kids come from wealthier families and wealthier families are able to offer additional support for their children to pursue colleges at any institution rather than intending to stay within state to save on tuition. Regardless of this endogeneity Regression Discontinuity Design (RDD) can be helpful in assessing the causal effects around the cutoff in which we'd expect individuals just on either side to be relatively similar.

$D_i = 1$ if $X_i \geq c$	**$D_i = 1$ if $X_i \geq c$**	**Scholarship if MCAS Score ≥ 75th percentile**
$D_i = 0$ if $X_i < c$	**$D_i = 0$ if $X_i < c$**	**No Scholarship if MACS Score < 75th percentile**



By comparing observations lying closely on either side of the threshold, it is possible to estimate the average treatment effect of scholarship eligibility on likelihood of enrollment into an in-state public college. Despite the lack of experimental design here, RDD can elicit causal effects by looking at the treatment groups (scholarship eligible) to the counterfactual outcome of non-Adam scholarship eligible students. The specific ATE estimate would be for the population of students with an MACS score at/equal to the 75th percentile (the cutoff), X = c. The average treatment effect would essentially be the difference in the likelihood of attending an in-state public college at the 75th percentile/cutoff.

b. (1 point) What assumptions are necessary for the design you described in (a) to yield an unbiased estimate of this effect? How might you test for the validity of these assumptions?

In order for RDD estimates to be unbiased, we need to assume the following:

1) **Continuity Assumption: $E[Y_1 | X]$ and $E[Y_0 | X]$ are continuous at X = c**
    a) **If this assumption holds true, the treatment effect is the difference of $E[Y_1 | X = c] - E[Y_0 | X = c]$ = E [Y | X just above cutoff] - $E[Y_0 | X$ just below cutoff]**
        i) **This assumption would be violated if there were differences between individuals just above or just below the cutoff, that are not explained by the treatment. I can't think of how this violation would apply in this scenario and I image it would be rather exhausting and difficult for students to try to intentionally manipulate the running variable of SAT score in order to gain or avoid treatment (harder than reporting or adjusting hours worked for welfare benefits). Maybe students could take the MACS multiple times if you were right below the threshold and knew of the cutoff. If students are allowed to take the MCAS multiple times, and there was a fee, which I'd expect, I would then suspect some difference between students right below and right above this threshold as students with more financial resources could retest until they achieved an adequate score to be eligible for the scholarship. Also students that would be first generation college students, would be less likely to have family members "in the know" about scholarship opportunity of this sort.**

Using the RDD you described above, researchers estimated the effect of Adams Scholarship eligibility on in-state public college enrollment and graduation rates. The scholarship winners included in the RDD study were 6.9 percentage points more likely to enroll in an in-state public college than the scholarship losers in the study (30.7 percent vs. 23.8 percent). However, both groups had approximately the same overall college enrollment rates; the increase in enrollment rates for in-state public colleges among the scholarship winners was offset by a decrease in enrollment rates for out-of-state and private colleges.

Surprisingly, the scholarship winners were 2.5 percentage points *less* likely to have graduated within 6 years of entering college than the scholarship losers. Critics of the program argue that this is because the scholarship induced students to pass on higher-quality schools (non-Adams schools) in favor of the lower-quality in-state public colleges covered by the scholarship (Adams schools).

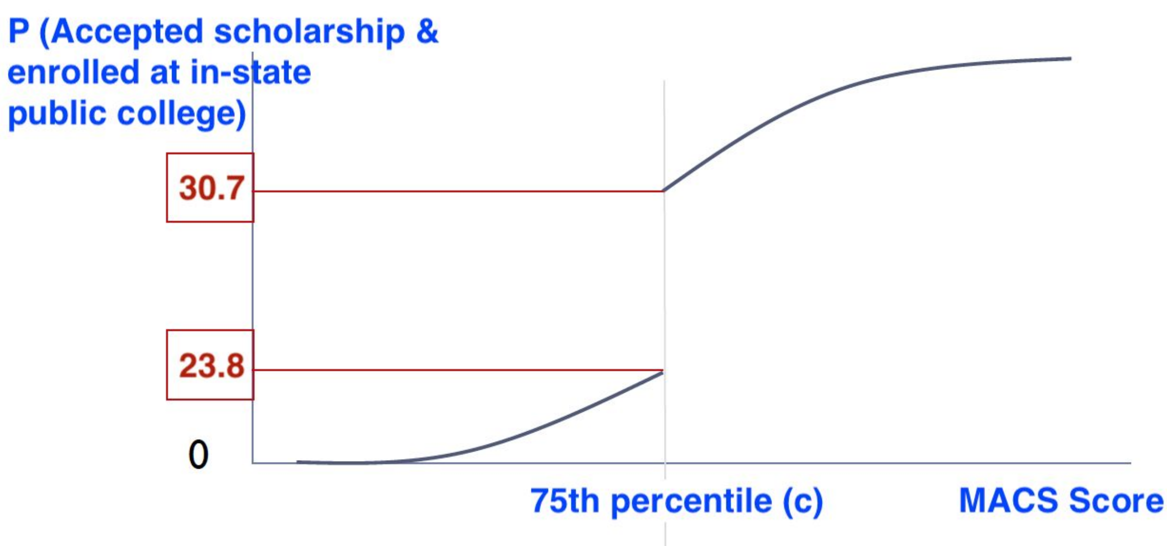
c. (*1 point*) Use a "fuzzy" RDD to estimate the effect of passing on a non-Adams school and enrolling in an Adams school on a student's likelihood of graduating within 6 years.

X = running variable MACS score
D = treatment of offered scholarship
c - cutoff at 75th percentile
Y - outcome is graduation within 6 years
Z - instrumental variable (likelihood of attending in-state public college)

"Fuzzy" RDD is used when the cutoff does not directly or exactly determine treatment. In this scenario, the cutoff of c at the 75th percentile of MACS score only indicates that a treatment of an Adam's Scholarship is offered, it does not mean that all students are immediately and automatically enrolled at an in-state public college upon scoring at or above the 75th percentile. A "fuzzy" RDD instead will create a discontinuity in probability of receiving treatment, the probability of enrolling in an in-state public college.

$Z_i = \{ \frac{1 \text{ if } Xi \geq c}{0 \text{ if } Xi < c}$

Through such we estimate the effect of treatment (scholarship offered) for compliers (accept scholarship and enroll at in-state public college) whose treatment D, depends on Z.



E [D | Z = 1] - E [D | Z = 0] = 30.7 - 23.8 = 6.9

The scholarship winners were 6.9 percentage points more likely to enroll in an in-state public college than those not offered or awarded a scholarship.

Let's hypothetically say we have have fractional units of people, and it was a cohort of 200 and 50 were awarded scholarships (scored at or above the 75th percentile), this means that about 15.35 scholarship winning students (30.7%) enrolled into an Adams school. This would mean that of the

remaining 150 scholarship losers, about 35.7 students enrolled into an Adams school. The extra information that overall college enrollments were fairly equal between these two groups doesn't impact calculations for this assignment.

We know that scholarship winners were 2.5 percentage points *less* likely to have graduated within 6 years of entering college than the scholarship "losers." So let's hypothetically say all (100%) scholarship losers graduated within 6 years and 97.5% of scholarship winners graduated within 6 years. So if we have 15.35 scholarship winners and 35.7 scholarship losers attending Adams schools, and there is a 2.5% difference in graduation within 6 years, and we assume 100% of scholarship losers graduated within 6 years = 35.7, and 97.5% of scholarship winners graduated within 6 years = 14.97.

$E [Y | Z = 1] - E [Y | Z = 0] = 2.5$

$LATE_{compliers} = E [Y_1 - Y_0 | D_1 > D_0] = \frac{E[Y|Z=1] - E[Y|Z=0]}{E[D|Z=1] - E[D|Z=0]} = \frac{2.5}{6.9} = 0.36$

d. (*1 point*) For what students does your estimate in (c) apply? Be precise.

With fuzzy RDD estimates and instrumental variables, we are estimating the Local Average Treatment Effect (LATE) for compliers, those who score at or above the 75th percentile, are offered the treatment and comply/ enroll in an in-state public college.

Question 2

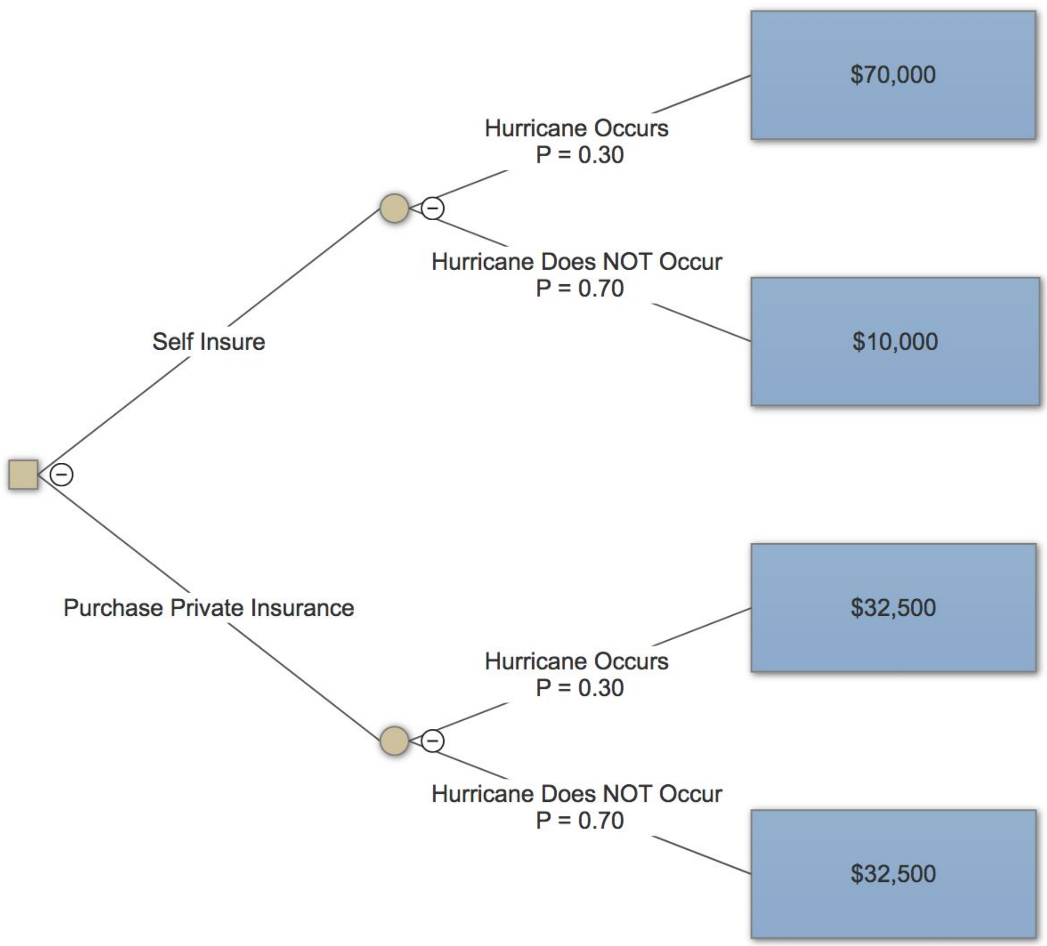
Suppose you are the Mayor of a Mississippi town on the coast of the Gulf of Mexico. Hurricane season is approaching, and you must decide whether or not to purchase disaster insurance for the two years remaining in your term. Each year there is a 30% chance of a serious hurricane. Your staff informs you that there are two options:

- The town can choose to self-insure for the next two years. This consists of purchasing disaster supplies and paying for repairs in the event of a hurricane. Each year, if a hurricane occurs, the cost to the town is $70k. Each year, if a hurricane doesn't occur, the cost to the town is $10k.
- The town can purchase disaster insurance from a private company. The private company then covers the cost of disaster supplies and any hurricane-related repairs. The private company offers a two-year insurance policy for $65k.

a. (*1 point*) Should you hire the private company or choose to self-insure? Draw a decision tree and solve it to determine which option minimizes expected cost.

As mayor I should decide to self-insure because my estimated self-insured costs are $9,000 less than if I were to purchase disaster insurance from a private company.

Purchase private insurance or self-insure ?



However the above decision tree only accounts for 1 year and we need to consider the combined probabilities that there are and are not severe hurricanes in year one and/or year two to compare the estimated costs and make a decision.

	Hurricane Year 1 Yes, P = 0.30 No, P = 0.70	Hurricane Year 2 Yes, P = 0.30 No, P = 0.70	Combined Probability & Costs for 2 years	Estimated Costs
Self-Insure	No, P = 0.70 $10,000	No, P = 0.70 $10,000	0.49 $20,000	$56,000 *(see calculation below table)*
	Yes, P = 0.30 $70,000	No, P = 0.70 $10,000	0.21 $80,000	
	No, P = 0.70 $10,000	Yes, P = 0.30 $70,000	0.21 $80,000	
	Yes, P = 0.30 $70,000	Yes, P = 0.30 $70,000	0.09 $140,000	
Private Insurance	No, P = 0.70 $32,500	No, P = 0.70 $32,500	0.49 $65,000	$65,000
	Yes, P = 0.30 $32,500	No, P = 0.70 $32,500	0.21 $65,000	
	No, P = 0.70 $32,500	Yes, P = 0.30 $32,500	0.21 $65,000	
	Yes, P = 0.30 $32,500	Yes, P = 0.30 $32,500	0.09 $65,000	

***($20,000 ✕ 0.49) + (0.21 ✕ $80,000) + (0.21 ✕ $80,000) + (0.09 ✕ $140,000) = $9,800 + $16,800 + $16,800 + $126,000 = $56,000**

b. (*1.5 points*) You are unsure about the cost to the town if a hurricane occurs (i.e. you are unsure about the $70k cost above). How high or low would the costs have to be to make you change your decision from part (a)?

For this question I plugged in my decision tree / table information into excel so I could manipulate the Yes / 0.30 / $70,000 cells to see how high or low the costs would have to be to change my decision from part (a) where I decided to self-insure because I would expect to save $9,000 in choosing such over the next two years. I first tried increasing the cost for a hurricane, when self-insured to $80,000 per year. But as shown below that resulted in an estimated cost of $62,000 which is not enough to change my mind.

	A	B	C	D	E
1		Year 1	Year 2	Combined Probability	Estimated Costs
2		No	No	No, No	
3		0.7	0.7	0.49	
4		$ 10,000	$ 10,000	$ 20,000	
5		Yes	No	Yes, No	
6		0.3	0.7	0.21	
7	Self-Insure	$ 80,000	$ 10,000	$ 90,000	$ 62,000
8		No	Yes	No, Yes	
9		0.7	0.3	0.21	
10		$ 10,000	$ 80,000	$ 90,000	
11		Yes	Yes	Yes, Yes	
12		0.3	0.3	0.09	
13		$ 80,000	$ 80,000	$ 160,000	
14		No	No	No, No	
15		0.7	0.7	0.49	
16		$ 32,500	$ 32,500	$ 65,000	
17		Yes	No	Yes, No	
18		0.3	0.7	0.21	
19	Private	$ 32,500	$ 32,500	$ 65,000	$ 65,000
20	Insurance	No	Yes	No, Yes	
21		0.7	0.3	0.21	
22		$ 32,500	$ 32,500	$ 65,000	
23		Yes	Yes	Yes, Yes	
24		0.3	0.3	0.09	
25		$ 32,500	$ 32,500	$ 65,000	

I then tried a cost of $85,000 per year and that resulted in an estimated cost equal to option two of purchasing private insurance. So ultimately, if the estimated costs for hurricane damage if self-insured per year were any amount more than $85,000, I would change my decision from part (a).

E2 fx =(D4*D3)+(D7*D6)+(D10*D9)+(D13*D12)

	A	B	C	D	E
1		Year 1	Year 2	Combined Probability	Estimated Costs
2		No	No	No, No	
3		0.7	0.7	0.49	
4		$ 10,000	$ 10,000	$ 20,000	
5		Yes	No	Yes, No	
6		0.3	0.7	0.21	
7	Self-Insure	$ 85,000	$ 10,000	$ 95,000	$ 65,000
8		No	Yes	No, Yes	
9		0.7	0.3	0.21	
10		$ 10,000	$ 85,000	$ 95,000	
11		Yes	Yes	Yes, Yes	
12		0.3	0.3	0.09	
13		$ 85,000	$ 85,000	$ 170,000	
14		No	No	No, No	
15		0.7	0.7	0.49	
16		$ 32,500	$ 32,500	$ 65,000	
17		Yes	No	Yes, No	
18		0.3	0.7	0.21	
19	Private	$ 32,500	$ 32,500	$ 65,000	$ 65,000
20	Insurance	No	Yes	No, Yes	
21		0.7	0.3	0.21	
22		$ 32,500	$ 32,500	$ 65,000	
23		Yes	Yes	Yes, Yes	
24		0.3	0.3	0.09	
25		$ 32,500	$ 32,500	$ 65,000	

I21 fx

	A	B	C	D	E
1		Year 1	Year 2	Combined Probability	Estimated Costs
2		No	No	No, No	
3		0.7	0.7	0.49	
4		$ 10,000	$ 10,000	$ 20,000	
5		Yes	No	Yes, No	
6		0.3	0.7	0.21	
7	Self-Insure	$ 85,001	$ 10,000	$ 95,001	$ 65,001
8		No	Yes	No, Yes	
9		0.7	0.3	0.21	
10		$ 10,000	$ 85,001	$ 95,001	
11		Yes	Yes	Yes, Yes	
12		0.3	0.3	0.09	
13		$ 85,001	$ 85,001	$ 170,002	
14		No	No	No, No	
15		0.7	0.7	0.49	
16		$ 32,500	$ 32,500	$ 65,000	
17		Yes	No	Yes, No	
18		0.3	0.7	0.21	
19	Private	$ 32,500	$ 32,500	$ 65,000	$ 65,000
20	Insurance	No	Yes	No, Yes	
21		0.7	0.3	0.21	
22		$ 32,500	$ 32,500	$ 65,000	
23		Yes	Yes	Yes, Yes	
24		0.3	0.3	0.09	
25		$ 32,500	$ 32,500	$ 65,000	

The town's budget is not large; the town has appropriated a maximum of $100k for hurricane-related costs over the entire two-year period. If the town spends more than $100k, it must raise taxes or cut other spending, so your political consultants recommend that you decrease your utility by $4 per dollar actually spent over $100k. (Thus, the utility of a $120k expenditure would be -$100k – 4*$20k = -$180k.)

c. (*1.5 points*) How does the expected cost of each option you calculated in (a) compare now with the expected utility? Would it change your preferred option?

Given this change in budget constraints, we must adjust the probabilistic cost in the scenario where we have a hurricane in year one and year two, because that is the only scenario in which we'd go over the $100k budget. Again in excel I adjusted this specific cell of $140,000 - originally the formula was just the added cost of a hurricane in year one and year two, but now we must build in added *penalty like/utility costs for the $40,000 over the $100k budget limit. Similar to the information provided in the prompt, with a decreased utility of $4 per dollar actually spent over $100k our formula for our $140k cost would read: -$100k - $4*$40k = $260k

The formula in excel was originally: C13+B13 = $140,000 if I change that cell to the new expected cost of $260,000 our total estimated would increase by $10,800, in which case my preferred option from part (a) would in fact change to instead purchase private insurance.

	A	B	C	D	E	F
1		Year 1	Year 2	Combined Probability	Estimated Costs	
2		No	No	No, No		
3		0.7	0.7	0.49		
4		$ 10,000	$ 10,000	$ 20,000		
5		Yes	No	Yes, No		Difference from
6		0.3	0.7	0.21		Part (a)
7		$ 70,000	$ 10,000	$ 80,000	$ 66,800	$ 10,800
8	Self-Insure	No	Yes	No, Yes		
9		0.7	0.3	0.21		
10		$ 10,000	$ 70,000	$ 80,000		
11		Yes	Yes	Yes, Yes		
12		0.3	0.3	0.09		
13		$ 70,000	$ 70,000	$ 260,000		
14		No	No	No, No		
15		0.7	0.7	0.49		
16		$ 32,500	$ 32,500	$ 65,000		
17		Yes	No	Yes, No		Savings for
18		0.3	0.7	0.21		Option 2
19	Private	$ 32,500	$ 32,500	$ 65,000	$ 65,000	$ 1,800
20	Insurance	No	Yes	No, Yes		
21		0.7	0.3	0.21		
22		$ 32,500	$ 32,500	$ 65,000		
23		Yes	Yes	Yes, Yes		
24		0.3	0.3	0.09		
25		$ 32,500	$ 32,500	$ 65,000		

Rather than a two-year policy, assume the private company offers policies that cover a single year for $32.5k. The town can make different decisions for each year (e.g. the town can purchase insurance in the first year and choose to self-insure in the second year). Moreover, in the second year, the town can make its decision after seeing whether or not a hurricane occurred in the first year. Continue to assume that, as in part (c), the town has only appropriated $100k for hurricane costs over the entire two-year period (and hence, the utility of a $120k expenditure would be -$100k – 4*$20k = -$180k).

d. (*2 points*) What decision(s) should the town make? Support your answer with a decision tree.

For the first year I'd self-insure because it is estimated to be the cheaper option, by $4,500.

<u>In the event that a hurricane did not actually happen in year 1</u>, I'll feel pretty good about this decision and the town would have only spent $10,000 ($22,500 less than what we would have spent had we purchased private insurance) and we'll have $90,000 remaining in our budget for year two. In this scenario I choose to self-insure again because it's again estimated to be cheaper than private insurance (though they're probably not, for this assignment we are assuming independence as in previous year hurricanes are independent of probability of severe hurricane in year 2). And I have

adequate room in my budget that should a we have a hurricane, which would cost us $70,000 totallying $80k, we'd still be under budget. Decision columns indicate "choice" / square nodes and year 1 and year 2 columns indicate "chance" / circle nodes.

Options Yr 1	Year 1	Year 1 Estimated Costs	Decision
Self-Insure	No 0.7 $ 10,000	$ 28,000	Self-Insure
	Yes 0.3 $ 70,000		
Purchase Private Insurance	No 0.7 $ 32,500	$ 32,500	
	Yes 0.3 $ 32,500		
	Counterfactual Saving	$ 4,500	
Budget	$ 100,000		

What Happened Yr 1?	Reflection	Options Yr 2	Year 2	Year 2 Estimatd Costs	Year 1 Actual + Year 2 Estimated	Decision	Actual Yr 2	Two Year Actual
No Hurricane $ 10,000	Remaining Budget $ 90,000.00 Counterfactual Saving $ 22,500 *Happy. Adequate in budget. Feel good about saving money.* *Will Choose to Self-Insure again*	Self-Insure	No 0.7 $ 10,000	$ 28,000	$ 38,000	Self-Insure	No Hurricane $ 10,000	$ 20,000
			Yes 0.3 $ 70,000				Yes Hurricane $ 70,000	$ 80,000
		Purchase Private Insurance	No 0.7 $ 32,500	$ 32,500				
			Yes 0.3 $ 32,500					

<u>In the event that a hurricane does actually happen in year 1</u>, we'll have spent $70,000 of the $100k budget and we'll have a counterfactual loss of $37,500 (not that it matters because it's a sunk cost). In this scenario, again you'd suspect people to look at estimated costs based on the probability of of a severe hurricane and again assume independence. That's all fine and dandy, <u>but</u> there's a lot more to consider now that we a utility penalty for money spent over $100k, and so we may not want to base a decision solely on probability estimates. In this scenario I can either self-insure again, at which point we have 70% of a no hurricane or a 30% of a severe hurricane, costing me either $80,000 or $140,00 respectively. Even though it's more likely we'll have no hurricane, once you consider the utility of $140k expenditure costing $260,000 I may be very very risk averse. And I'd rather choose to purchase private insurance at $32,500 with 100% certainty and have a total utility expenditure of $110,00 rather than risk $260,000 even with a 30% probability.

What Happened Yr 1?	Reflection	Options Yr 2	Year 2	Year 2 Estimatd Costs	Year 1 Actual + Year 2 Estimated	Decision	Actual Yr 2	Two Year Act	Plus Utility
Hurricane $ 70,000	Remaining Budget $ 30,000.00 Counterfactual Saving $ (37,500) *Not happy. Spent most of budget in year 1.* *Even if Self insurance offers a better estimated cost* *I will be risk averse because worst case scenario is horrible*	Self-Insure	No 0.7 $ 10,000	$ 28,000	$ 98,000	Self-Insure	No Hurricane $ 10,000	$ 80,000	
			Yes 0.3 $ 70,000				Yes Hurricane $ 70,000	$ 140,000	$ 260,000
		Purchase Private Insurance	No 0.7 $ 32,500	$ 32,500	$ 102,500	Purchase Private Insurance	No Hurricane $ 32,500	$ 102,500	$ 110,000.00
			Yes 0.3 $ 32,500				Yes Hurricane $ 32,500		