Problem Set 2

Last time we talked about the difficulty to satisfy any version of **A3** with non-experimental data. Now we discuss how even with experimental data it is difficult to estimate some well-defined treatment effects **in the presence of heterogenous treatment effects**. This problem set is intended to shape your understanding of different treatment effects of interest (ATE, TOT, LATE, ITT) and how we can learn (or not) about some of them from experimental data. As well we will discuss what happens if there is non-perfect compliance with the experimental treatment assignment.

Question 1 - Issues of Compliance

A) **Question:** *In the example of the MDVE, the treatment effect differed for a variety of reasons including that police simply forgot to bring the color-coded note pad. If this was the only form of bias and you would run a simple OLS regression of the outcome on treatment, why would this be an unbiased estimate of the ATE? Show this algebraically and then interpret your results.*

Answer: First note that the OLS estimate with a constant gives the same coefficient as the OLS estimate when using the data in mean-deviations form and skipping the constant coefficient. Let us do the later, for notational simplicity. Let α_i be the true treatment effect for individual i, $\bar{\alpha}$ is the ATE, y_i is the outcome variable (here: re-offence rates) and R_i is a dummy being one if the individual received treatment. Then

$$y_i = \alpha_i R_i + \epsilon_i = \bar{\alpha} R_i + \underbrace{\epsilon_i + (\alpha_i - \bar{\alpha}) R_i}_{\epsilon_i^*} \quad (1)$$

$$\hat{\alpha}_{\text{OLS}} = \bar{\alpha} + (R'R)^{-1} R' \epsilon^* = \bar{\alpha} + \frac{\sum_{R_i=1} \epsilon_i}{T_R} + \frac{\sum_{R_i=1} (\alpha_i - \bar{\alpha})}{T_R} \quad (2)$$

$$E[\hat{\alpha}_{\text{OLS}}] = \bar{\alpha} + E[\epsilon_i | R_i = 1] + E[\alpha_i - \bar{\alpha} | R_i = 1] \quad (3)$$

(1) is the true relation, (2) is the OLS estimate and (3) is its expectation. This should be $\bar{\alpha}$ for OLS to be an unbiased estimate of ATE.

Conclusion: ϵ_i **and** the size of the treatment effect need to be mean independent of receiving treatment. With random assignment and perfect compliance this will be the case. However, generally there won't be perfect compliance. But then - even with random assignment - the OLS estimate is not unbiased for ATE. This is because then $E[\alpha_i - \bar{\alpha} | R_i = 1]$ is probably not 0 since those with and over-average treatment effect ($\alpha_i > \bar{\alpha}$) select into treatment. [So among those with $R_i = 1$ the average of $\alpha_i - \bar{\alpha}$ is bigger than 0.]

B) **Question:** *In the experiment we investigate, the treatment assignment was random. But was there perfect compliance? Determine the compliance rates using the data you were provided. You can do this by using the command:* `tab t_random t_final, row`

Answer:

'Coddle' is defined as 'treatment' here (though that seems a little counter-intuitive). Hence we have few **'always takers'**, but quite some **'never takers'**.

C) **Question:** *The researchers in the MDVE kept track of why they may have given a different treatment than was assigned. To see these reasons, look at the variable reason2. You can do this by typing:* `tab reason2`

Answer:

We see that there are quite some 'never-takers'. For them the police probably decides not to give the assigned 'coddle' treatment since they are hard-core guys for whom the effect of this comparatively soft treatment on the re-offence rates would be particularly high. What does this


```
. tab t_random t_final, row
```

Key
frequency
row percentage

Police Sheet Color	Final Disposition arrest	advise	suspect t	other	Total
pink	91 97.85	0 0.00	1 1.08	1 1.08	93 100.00
yellow	19 17.27	84 76.36	5 4.55	2 1.82	110 100.00
blue	26 20.47	5 3.94	83 65.35	13 10.24	127 100.00
Total	136 41.21	89 26.97	89 26.97	16 4.85	330 100.00

```
. tab reason2
```

Reason for Not Complying with Random Assignment	Freq.	Percent	Cum.
blank	313	94.85	94.85
party assaults police officer	1	0.30	95.15
victim makes citizen's arrest	1	0.30	95.45
injury constitutes an aggravated assaul	5	1.52	96.97
victim has order of protection against/	1	0.30	97.27
other	4	1.21	98.48
unknown	5	1.52	100.00
Total	330	100.00	

mean for $E[\alpha_i - \bar{\alpha}|R = 1]$? It would be < 0 since only those who are comparatively not so dangerous get the treatment and hence the bias of OLS as estimate of ATE is downwards.

Question 2 - Measuring treatment effects

In class we discussed the Angrist (2006) paper which analyzes the Minnesota Domestic Violence Experiment (MDVE) in a Instrumental Variable framework. Define the assigned treatment of coddling as T_i and the received treatment of coddling at R_i. Also recall our counterfactuals: Y_{1i} is the outcome for individual i if they received treatment. Y_{0i} is the outcome for individual i if they did not received treatment. In practice we only observe Y_{1i} for the treatment group and Y_{0i} for the control group so that for any individual i, we don't know the counterfactual unobserved outcome.
While in general we want to estimate the average treatment effect, $ATE = E[Y_{1i} - Y_{0i}]$, sometimes compliance is an issue. As a result, instead of ATE, experimental studies may focus on the intent to treat effect, $ITT = E[Y_{1i}|T_i = 1] - E[Y_{0i}|T_i = 0]$. This is because looking only at what treatment was actually delivered, the treatment on the treated or $TOT = E[Y_{1i}|R_i = 1] - E[Y_{0i}|R_i = 0]$, cannot be estimated consistently/unbiasedly since some individuals selected themselves into the control and treatment group. Hence the group of non-treated will not be a good control for the group of treated individuals. We discussed why in this context, an estimable treatment effect may be the local average treatment effect or $LATE = E[Y_{1i} - Y_{0i}|R_{1i} > R_{0i}]$. R_i indicates whether individual i received treatment, R_{1i} indicates whether individual i when assigned treatment would receive treatment and R_{0i} indicates whether individual i when not assigned treatment would receive treatment. $R_{1i} = R_{0i} = 1$

are 'always-takers' and $R_{1i} = R_{0i} = 0$ *are 'never-takers'. In all of this assume there are no 'defiers' (monotonicity assumption)*[1]*.*

A) **Question:** *Show that in general, with non-compliance, ITT will be smaller than the true ATE. What is the intuition behind this? (HINT: Think about when ITT and ATE will be equal.)*

Answer: Consider the definitions of both treatment effects.

$$\text{ATE} \equiv E[Y_{1i} - Y_{0i}]$$

is the expected effect of treatment on a randomly drawn person from the population (Wooldrige, p.604). Let $Y_i \equiv Y_{0i}(1 - R_i) + Y_{1i}R_i$ be the observed outcome for individual i.

$$\text{ITT} \equiv E[Y_i|T_i = 1] - E[Y_i|T_i = 0]$$

is expected difference in the outcomes of those assigned to treatment and those who were not assigned to treatment. Also,

$p_a \equiv Pr[R = 1|T = 0]$ is the proportion of always takers and

$p_n \equiv Pr[R = 0|T = 1]$ is the proportion of never takers.

Then (using the law of iterated expectations)

$$\begin{aligned}
ITT &= E[Y_i|T = 1] - E[Y_i|T = 0] \\
&= E[Y_i|T = 1, R = 1] \cdot (1 - p_n) + E[Y_i|T = 1, R = 0] \cdot (p_n) \\
&\quad -E[Y_i|T = 0, R = 1] \cdot (p_a) + E[Y_i|T = 0, R = 0] \cdot (1 - p_a) \\
&= E[Y_i|T = 1\&R = 1] - E[Y_i|T = 0\&R = 0] \\
&\quad -p_n(E[Y_i|T = 1, R = 1] - E[Y_i|T = 1, R = 0)) \\
&\quad -p_a(E[Y_i|T = 0, R = 1] - E[Y_i|T = 0, R = 0))
\end{aligned}$$

Under **perfect compliance** with the treatment assignment $p_n = p_a = 0$ and $E[\bullet|T = 1, R = 1] = E[\bullet|T = 1]$ and hence[2]

$$\text{ITT} = E[Y_{1i}|T = 1] - E[Y_{0i}|T = 0]$$

and under **randomization** (Y_{1i}, Y_{0i}) is fully independent of T. Hence

$$\text{ITT} = E[Y_{1i} - Y_{0i}] = \text{ATE}$$

Conclusion: With randomization of the treatment assignment and perfect compliance ITT is the same as ATE. However without perfect compliance ITT is generally smaller than ATE.[3]

[1] 'Defiers' are those who, when assigned treatment would not take it and when not assigned treatment, would take it anyways.

[2] Note that we replace Y_i with what it actually is.

[3] You might note that one way for this to hold is that both among those who are assigned treatment on those who are not assigned treatment, the group who then actually receives the treatment (i.e. R=1) has higher outcomes than the group who did not (i.e. R=0). Then the terms in brackets are positive and hence the bias is negative.

B) **Question:** *In general, TOT and LATE will not be the same. This is because TOT is a weighted average of two effects: one on always-takers and one on compliers. Show that this is the case.*

Answer: Again let us recap the treatment effects:

$$\text{TOT} \equiv E[Y_{1i} - Y_{0i} | R_i = 1]$$

"is the mean effect for those who actually participated in the program" (Wooldrige, p.605). LATE is calculated pretty much like the ATE, but only using compliers:

$$\text{LATE} \equiv E[Y_{1i} - Y_{0i} | R_{1i} > R_{0i}]$$

In other words, it is "the average treatment effect for those who would be induced to participate by changing T from zero to one" (Wooldridge, p.635). Now, first note that there are two types who receive treatment ($R_i = 1$), compliers (for whom $R_{1i} > R_{0i}$) and always-takers (for whom $R_{1i} = R_{0i} = 1$). Then (using the law of iterated expectations)

$$\begin{aligned} \text{TOT} &\equiv E[Y_{1i} - Y_{0i} | R_i = 1] \\ &= \underbrace{E[Y_{1i} - Y_{0i} | R_{1i} > R_{0i}]}_{\text{treatment effect of compliers}} \cdot Pr(R_{1i} > R_{0i} | R_i = 1) \\ &+ \underbrace{E[Y_{1i} - Y_{0i} | R_{1i} = R_{0i} = 1]}_{\text{treatment effect of always-takers}} \cdot Pr(R_{1i} = R_{0i} = 1 | R_i = 1) \end{aligned}$$

Conclusion: TOT is a weighted average between compliers and always takers. When there are no 'always-takers' ($Pr(R_{1i} = R_{0i} = 1 | R_i = 1) = 0$) then TOT is equal to LATE. This is very useful, since an IV estimate of the treatment effect will be a consistent estimate of LATE - and if there are no 'always-takers' hence TOT.

C) **Question:** *In question 1 you showed that in the MDVE, there was mostly only one-sided non-compliance. If this was true, how does LATE relate to TOT in this case?*

Answer: The IV estimator will consistently estimate LATE. But we have very few 'always takers', almost none. Then LATE = TOT. Hence in our case the IV estimator will consistently estimate TOT.

Question 3 - Replicating Angrist's Results

To get used to working with your data and interpreting Stata output, please replicate the results found in Angrist. Don't worry if your estimates are not exactly the same as those presented in the paper (they won't be for complicated reasons related to simulated outcomes that are done slightly differently in this data than in the data in the paper).

A) **Question:** *Begin with the 'Reduced Form' estimate of the probability an individual re-offends on the assigned treatment. You can do this with the command:*
`regress reoffend1 coddle_assigned`
What is your estimate? Explain why this estimate can be interpreted as the ITT.

Answer:

We estimate that the treatment 'coddling' (rather than arresting) increases in the re-offence rate by 0.10. The OLS regression simply gives us the difference of the re-offence rates of those who were assigned the 'coddle' treatment and those who were not. This is the sample analog of the ITT since it uses assigned treatment, regardless of actual receipt of the treatment to measure the treatment effect.

```
. regress  reoffend1  coddle_assigned

      Source |       SS       df       MS              Number of obs =     330
-------------+------------------------------           F(  1,   328) =    4.85
       Model |  .714701115     1  .714701115           Prob > F      =  0.0284
    Residual |  48.376208    328  .147488439           R-squared     =  0.0146
-------------+------------------------------           Adj R-squared =  0.0116
       Total |  49.0909091   329   .14921249           Root MSE      =  .38404

------------------------------------------------------------------------------
   reoffend1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
coddle_ass~d |   .1034436   .0469916     2.20   0.028     .0110006    .1958865
       _cons |   .1075269   .0398233     2.70   0.007     .0291855    .1858682
------------------------------------------------------------------------------
```

B) **Question:** *Now add some covariates. You can do this by the command:*
`regress reoffend1 coddle_assigned  y82 q1 q2 q3 nonwhite mixed anyweapon s_influence`.
Does your point estimate change much? Why not? What happens to the R-Squared? Why is this important?

Answer:

```
. regress reoffend1 coddle_assigned  y82 q1 q2 q3 nonwhite mixed anyweapon s_influence

      Source |       SS       df       MS              Number of obs =     330
-------------+------------------------------           F(  9,   320) =    1.31
       Model |  1.7384865     9  .193165167           Prob > F      =  0.2330
    Residual |  47.3524226   320  .147976321           R-squared     =  0.0354
-------------+------------------------------           Adj R-squared =  0.0083
       Total |  49.0909091   329   .14921249           Root MSE      =  .38468

------------------------------------------------------------------------------
   reoffend1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
coddle_ass~d |     .09811   .0475365     2.06   0.040     .0045865    .1916336
         y82 |   .0132675   .0529572     0.25   0.802    -.0909207    .1174558
          q1 |  -.0368912   .0805851    -0.46   0.647    -.1954347    .1216524
          q2 |   .0529932   .0651508     0.81   0.417    -.0751848    .1811712
          q3 |  -.0292877   .0698676    -0.42   0.675    -.1667456    .1081701
    nonwhite |   .0276911   .0436444     0.63   0.526     -.058175    .1135572
       mixed |  -.0198068   .0496336    -0.40   0.690    -.1174563    .0778427
   anyweapon |  -.0770743   .0489911    -1.57   0.117    -.1734597     .019311
 s_influence |   .0169292   .0439631     0.39   0.700     -.069564    .1034224
       _cons |   .1049738   .0768676     1.37   0.173     -.046256    .2562035
------------------------------------------------------------------------------
```

The point estimate does not change much. Because assignment was random, the point estimate should not change when including control variables. The R-squared increased. This allows greater precision for the estimates because the residual variance is reduced.

C) **Question:** *Estimate the OLS treatment effect. You can do this with the command:*
`regress reoffend1 coddle_received`.
How does this compare to the estimates in A? Is this the true treatment effect? Why or Why not? [NOTE: You are using a different outcome variable here (reoffend2) because issues related to the outcome simulation. Don't worry about that and just pretend as if this is the same variable as in A and D]

Answer:

This is a little bit higher than the effect in 3.A (0.003 difference). If there were no compliance problems or selection bias for treatment versus control this would be an unbiased estimate of the average treatment effect (see question 1.A). Because there are compliance problems, we cannot expect this to be a good estimate of the average treatment effect.

D) **Question:** *Estimate the IV treatment effect. You can do this with the command:*
`ivreg  reoffend1 (coddle_received = coddle_assigned)`.
What type of treatment effect does the instrumental variables approach recover? Why?

Answer:

```
. regress reoffend2  coddle_received

      Source |       SS       df       MS              Number of obs =     330
-------------+------------------------------           F(  1,   328) =    6.39
       Model |  .901740724     1  .901740724           Prob > F      =  0.0119
    Residual |  46.2528047   328  .141014649           R-squared     =  0.0191
-------------+------------------------------           Adj R-squared =  0.0161
       Total |  47.1545455   329  .143326886           Root MSE      =  .37552

------------------------------------------------------------------------------
   reoffend2 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
coddle_rec~d |   .1062007    .041997     2.53   0.012     .0235832    .1888183
       _cons |   .1102941   .0322005     3.43   0.001     .0469486    .1736397
------------------------------------------------------------------------------
```

```
. ivreg  reoffend1  (coddle_received = coddle_assigned)

Instrumental variables (2SLS) regression

      Source |       SS       df       MS              Number of obs =     330
-------------+------------------------------           F(  1,   328) =    4.76
       Model | -.135460603     1 -.135460603           Prob > F      =  0.0298
    Residual |  49.2263697   328  .150080395           R-squared     =       .
-------------+------------------------------           Adj R-squared =       .
       Total |  49.0909091   329   .14921249           Root MSE      =   .3874

------------------------------------------------------------------------------
   reoffend1 |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
coddle_rec~d |   .1311702   .0601084     2.18   0.030     .0129236    .2494167
       _cons |    .104706   .0412729     2.54   0.012      .023513     .185899
------------------------------------------------------------------------------
Instrumented:  coddle_received
Instruments:   coddle_assigned
------------------------------------------------------------------------------
```

This recovers the LATE treatment effect. The point estimate is substantially higher than in the previous parts of this question.

E) **Question:** *Compute the mean re-offense rate. You can do this with the command:* `sum reoffend1` *From this compute the percent change in recidivism rates for the various estimates ITT, OLS and LATE presented in class. Is the relationship between the estimates as you predicted? Why or why not?*

Answer:

```
. sum reoffend1

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
   reoffend1 |       330    .1818182    .3862803          0          1
```

ITT The effect of 0.103 means an increase in the re-offence rate of 57%.

OLS The effect of 0.106 means an increase in the re-offence rate of 59%.

IV The effect of 0.131 means an increase in the re-offence rate of 73%.

In the light of question 2 we can make sense of these results. First remember that (under the assumption of no defiers) ATE is a weighted average of the average treatment effects on always-takers, never-takers and compliers and TOT is a weighted average of the average treatment effect on always takers and compliers. Since there no always-takers in our example, TOT is the average treatment effect of compliers and ATE is the average treatment effect of complieres and never-takers. Hence we would expect ATE to be bigger than TOT. We won't be able to find out what ATE is, but let us see which of the above estimates comes closest to estimate TOT? Remember what we said about the biases of OLS, ITT and IV:

- We know that here those who select out of treatment were those who have high re-offence rates anyways, hence a simple OLS estimate comparing the outcomes of those who got treated with those who did not get treated would underestimate TOT.

- We expect ITT to be less than TOT again because of non-compliance. If all those actually assigned the ‘coddle’ treatment would get it (and the other way around, though that does not seem to be a problem here), ITT would estimate TOT (and ATE) unbiasedly. However, here the worst (those who would have the highest re-offence rates) are actually taken out of the treated group and put into the non-treated group. This makes the treated group look rather favorable and hence our estimate of the treatment effect being biased downwards.
- We know that (under the monotonicity assumption, which seems reasonable here) IV will consistently estimate LATE. But since here there are few/no ‘always-takers’, this is the same as TOT.

"The coddle treatment, when applied, increases the re-offence rates by 73%" is hence our best answer!

Question 4 - Interpreting the results

The MDVE found convincing evidence that individuals who are arrested after they commit domestic violence are less likely to re-offend after arrest. If we manage to design the experiment carefully and tackle convincingly the issues in questions 1, 2 and 3 we say that the estimates are ***internally valid****. Many advocates and policy makers were concerned about police not arresting frequently enough and hence many states passed so called "Mandatory Arrest Laws" which required the police to arrest an offender when a domestic violence incident was reported. In my paper (Iyengar, 2009), I show that in states that passed these laws, domestic violence actually went up after the laws were passed. This exercise helps understand why experimental results* ***may not*** *be translated correctly into public policy. This second question, whether the results from the experiment translate to a possible policy, is a question of* ***external validity****. To answer it think about what is different when actually implementing the policy vs. running the experiment.*

A) **Question:** *Consider first the initial experiment: Someone reported a crime and then a police unit dispatched would apply a randomly assigned treatment. This meant that the treatment effect measured Pr(Reoffend|Arrest & Report). Does a law which mandates the police arrest replicate this experimental setting? Why or why not.*

Answer: A mandatory arrest laws increases the Pr(Arrest) while the experiment increased the Pr(Arrest|Report). Thus the effect of a mandatory arrest law is to estimate Pr(Reoffend|Arrest) unconditional on reporting.

As a policy maker we would be interested in decreasing Pr(Reoffend)! We can rewrite

$$\begin{aligned} Pr(\text{Reoffend}) &= Pr(\text{Reoffend}|\text{Report}) \cdot Pr(\text{Report}) + Pr(\text{Reoffend}|\text{No report}) \cdot Pr(\text{No report}) \\ &= Pr(\text{Reoffend}|\text{Arrest \& Report}) \cdot Pr(\text{Arrest}|\text{Report}) \cdot Pr(\text{Report}) \\ &+ Pr(\text{Reoffend}|\text{No arrest \& Report}) \cdot Pr(\text{No arrest}|\text{Report}) \cdot Pr(\text{Report}) \\ &+ Pr(\text{Reoffend}|\text{No report}) \cdot Pr(\text{No report}) \end{aligned}$$

The experiment showed that Pr(Reoffend|Arrest & Report) < Pr(Reoffend|No arrest & Report). Hence you might be tempted to pass a mandatory arrest law (increasing Pr(Arrest|Report)) in order to decrease the re-offence rates. However, while in the experiment Pr(Report) remained unchanged, with a mandatory arrest policy this might decrease. And since Pr(Reoffend|No report) is likely to be high, the re-offence rates (observed or not) might go up.

B) **Question:** *Iyengar (2009) uses a 'natural experiment' to measure the causal effect of mandatory arrest laws on domestic violence. The variation comes from the fact that some states passed mandatory arrest laws (treatment group) and some states did not (control group) and thus some individuals were 'as if' randomly assigned to treatment. What assumption is necessary for this* *to be a 'quasi' experiment? [HINT: Think about what the source of variation is and how that is related to unobserved factors.] What evidence could Iyengar provide to support this assumption?*

Answer: You need to think that states with mandatory arrest laws and states without them looked the same before the laws were passed. Show that variables related to arrest, reoffense rates, and violence levels are on average the same

C) **Question:** *Figure 1 above shows that while intimate partner violence went up in states with mandatory arrest laws, family violence (i.e. child abuse) went down after the laws were passed. The paper also notes that while intimate partner violence is most often reported by the victim, child abuse is typically reported by outside third parties (like doctors or teachers). Why does this help explain why the results in family violence more closely mirror those of the experiment while the results in intimate partner violence do not?*

Answer: Because victims do not report in the case of family violence, their reporting may not be reduced by the increased probability of arrest.

