MQE: Economic Inference from Data: Module 4: Randomized Control Trials

Claire Duquennois

6/9/2020

```r
library(knitr)
hook_output = knit_hooks$get('output')
knit_hooks$set(output = function(x, options) {
  # this hook is used only when the linewidth option is no
  if (!is.null(n <- options$linewidth)) {
    x = knitr:::split_lines(x)
    # any lines wider than n should be wrapped
    if (any(nchar(x) > n)) x = strwrap(x, width = n)
    x = paste(x, collapse = '\n')
  }
  hook_output(x, options)
})
```

Module 4: Randomized Control Trials

The Experimental Ideal:

Getting causal effects is HARD!

The Experimental Ideal:

Getting causal effects is HARD!

So where do we go from here?

Randomized control trials (RCT's) aka the "Gold standard" of experimental designs

Random assignment and the selection problem

The idea: use random assignment remove selection bias

Suppose a constant treatment effect: $Y_i(1) - Y_i(0) = \tau$, a constant. For observation i we have that

$$Y_i = Y_i(0) + \tau D_i$$
$$Y_i = E[Y_i(0)] + \tau D_i + Y_i(0) - E[Y_i(0)]$$
$$Y_i = \alpha + \tau D_i + \eta_i$$

where $\alpha = E[Y_i(0)]$, $\tau = Y_i(1) - Y_i(0)$, and η_i is the random part of $Y_i(0)$ since $\eta_i = Y_i(0) - E[Y_i(0)]$.

Random assignment and the selection problem

The expected outcomes for someone with treatment ($D_i = 1$), and without treatment ($D_i = 0$) are given by

$$E[Y_i(1)] = \alpha + \tau + E[\eta_i | D_i = 1]$$
$$E[Y_i(0)] = \alpha + E[\eta_i | D_i = 0]$$

so that we can break down the difference between these outcomes as

$$E[Y_i(1)] - E[Y_i(0)] = \underbrace{\tau}_{\text{treatment effect}} + \underbrace{E[\eta_i | D_i = 1] - E[\eta_i | D_i = 0]}_{\text{selection bias}}.$$

Random assignment and the selection problem

Selection bias will bias our estimate of τ if those who select into treatment have a different expected outcome compared to those who do not select into treatment:

$$E[Y_i(0)|D_i = 1] \neq E[Y_i(0)|D_i = 0].$$

This is because treatment is not random: $\{Y_i(1), Y_i(0)\} \not\perp D_i$.

There is no reason to believe that those who select into treatment have the same expected outcome as those who do not, if they were to be treated, that is to say, it is possible (and even likely) that

$$\underbrace{E[Y_i(0)|D_i = 0]}_{\text{observed}} \neq \underbrace{E[Y_i(0)|D_i = 1]}_{\text{unobserved}} \neq E[Y_i(0)]$$

Random assignment and the selection problem

The conditional independence assumption allows us to control for selection bias by conditioning on **observed characteristics**...

... **unobserved characteristics** that we cannot control for will often also bias our estimates.

Random assignment solves all of these selection bias problems.

Random assignment and the selection problem

Random assignment makes D_i independent of potential outcomes:

$$\{Y_i(1), Y_i(0)\} \perp D_i.$$

With random assignment, we know that in expectation,

$$\underbrace{E[Y_i(0)|D_i = 0]}_{\text{observed}} = \underbrace{E[Y_i(0)|D_i = 1]}_{\text{unobserved}} = E[Y_i(0)]$$

and

$$\underbrace{E[Y_i(1)|D_i = 0]}_{\text{unobserved}} = \underbrace{E[Y_i(1)|D_i = 1]}_{\text{observed}} = E[Y_i(1)]$$

Random assignment and the selection problem

Thus, the causal **Average Treatment Effect (ATE)**, $\bar{\tau}$, is

$$\bar{\tau} = E[Y_i(1)] - E[Y_i(0)] = \underbrace{E[Y_i(1)|D_i = 1]}_{\text{observed}} - \underbrace{E[Y_i(0)|D_i = 0]}_{\text{observed}}$$

$$= E[Y_i|D_i = 1] - E[Y_i|D_i = 0].$$

and we can easily estimate $\bar{\tau}$, by taking the difference between the average value of Y_i in the treatment group and the average value of Y_i in the control group.

RCT estimation

RCT regressions are about as straigtforward as it gets.

As modeled above, you can estimate

$$Y_i = \alpha + \tau D_i + \eta_i$$

where $\alpha = E[Y_i(0)]$, $\tau = Y_i(1) - Y_i(0)$, and η_i is the random error term.

RCT estimation

The treatment effect will be given by

$$E[Y_i(1)] - E[Y_i(0)] = \underbrace{\tau}_{\text{treatment effect}} + \underbrace{E[\eta_i|D_i = 1] - E[\eta_i|D_i = 0]}_{\text{selection bias}}.$$

With proper randomizing,

$$E[\eta_i|D_i = 1] = E[\eta_i|D_i = 0]$$

so there is no selection bias giving us an unbiased estimate of τ.

RCT Simulation

I am a principal of a large school and I want to evaluate how access to small reading groups with a paraprofessional helps improve 4th grade test scores.

I take all the 4th graders and randomly assign 30 percent of them to treatment (participating in the reading groups) with therest to the control group which continued with class as normal.

RCT Simulation

I generate a set of simualted data:

```
set.seed(1999)

scores5<-as.data.frame(rep(c(1,2,3,4,5,6,7,8,9,10),times=30))
names(scores5)<-c("class")
scores5 <- fastDummies::dummy_cols(scores5, select_columns = "class")
scores5$error<-rnorm(300, mean=0, sd=10)

#treatment indicator
scores5$treat<-rbinom(300,1,0.3)

#mean reading score
alpha=75

#treatment effect
tau=10

#the data generating process: notice the class does affect a student's score
scores5$read4<-alpha+tau*scores5$treat+scores5$error+4*scores5$class_1+(-6)*scores5$class_2+8*scores5$clas
```

RCT Simulation

```
rct1<-felm(read4~treat,scores5)
stargazer(rct1, type="latex", header=FALSE)
```

Table 1

	Dependent variable:
	read4
treat	11.229***
	(1.442)
Constant	77.150***
	(0.750)
Observations	300
R^2	0.169
Adjusted R^2	0.166
Residual Std. Error	11.092 (df = 298)
Note:	*p<0.1; **p<0.05; ***p<0.01

Because treatment was randomized, even though the class the student is in does affect their score, we recover an unbiased estimate of the true treatment effect ($\tau = 10$).

RCT Key assumption

The key assumption is that

$$E[\eta_i|D_i = 1] = E[\eta_i|D_i = 0] = 0.$$

▶ we cannot test this assumption directly

▶ BUT we can do a **balance test**: we check to see if observable characteristics among treatment and control groups are the same on average.

RCT Balance tests

- can be presented as a table of the following regressions $X_i = \beta_0 + \beta_1 D_i + \epsilon_i$ where X_i is a vector of characteristics being tested.

- are often presented as simple t-test tables testing the difference in means between the treatment and control groups.

RCT Balance Tests

Balance test variables should be

- ▶ characteristics at baseline, prior to treatment,
- ▶ or characteristics that would be unaffected by treatment.

Balance tests are often run on many variables:

- ▶ some may come up with a statistically significant difference by simple random chance
- ▶ if many are significantly different, this is a red flag that the key assumption does not hold
- ▶ There are corrections that can be implemented if the unbalanced variables are are of particular concern (look up Bonferroni correction)

RCT Simulation

Suppose the principal is concerned that there were some problems with the randomization.

She has access to some additional data. She adds it to her data set and does a balance test.

RCT Simulation

```r
#simulating covariates

#third grade test scores. Notice I am genereteing simulated academic scores that have a correlation to thei
scores5$read3<-alpha+scores5$error+rnorm(300,3,2)
scores5$math3<-alpha+scores5$error+rnorm(300,15,2)
scores5$hist3<-alpha+scores5$error+rnorm(300,5,2)
scores5$pe3<-rnorm(300,90,2)

#other 4th grade test scores: notice I am generating scores that correlated with their subject performance
scores5$hist4<-4*scores5$treat+scores5$hist3+rnorm(300,-2,2)
scores5$pe4<-scores5$pe3+rnorm(300,0,5)
scores5$math4<-2*scores5$treat+scores5$math3+rnorm(300,-5,3)

#student characteristics
scores5$female<-rbinom(300,1,0.5)
scores5$age<-runif(300,9,10)
scores5$height<-rnorm(300,1.3,0.2)

scoresmini<-scores5[,c("treat", "read4", "read3", "math3","hist3","pe3","hist4","pe4","math4","female", "a
knitr::kable(head(scoresmini))
```

treat	read4	read3	math3	hist3	pe3	hist4	pe4	math4	fema
0	86.32672	83.27039	96.31086	87.68543	91.03541	88.16632	93.95379	94.98303	
0	68.62170	79.22211	88.71313	79.91069	93.06738	76.14878	94.28598	81.82348	
1	105.03009	87.89982	102.82710	92.07489	86.74861	95.05262	82.04735	102.39419	
0	85.69802	94.08272	101.67523	95.73204	88.38979	94.73689	76.42173	97.08292	
0	83.33690	74.47617	90.96568	80.65065	93.68531	77.05651	86.33431	84.26472	
0	78.19827	80.55488	96.59311	84.92801	87.12997	83.51559	93.56600	92.40844	

RCT Simulation

```r
#as you can see, we have simulated some complex interrelationships between theses variables.
cor(scoresmini)
```

```
##             treat        read4       read3        math3       hist3
## treat    1.000000000  0.411103370  0.06210422  0.043810660  0.048985855
## read4    0.411103370  1.000000000  0.76370646  0.756595945  0.767164090
## read3    0.062104220  0.763706465  1.00000000  0.954947392  0.957691240
## math3    0.043810660  0.756595945  0.95494739  1.000000000  0.951163232
## hist3    0.048985855  0.767164090  0.95769124  0.951163232  1.000000000
## pe3     -0.134794987 -0.078928737  0.01977274  0.025802541  0.002345894
## hist4    0.210104738  0.799439890  0.92717436  0.910566252  0.965319194
## pe4     -0.043975444  0.001522679  0.06059868  0.067988360  0.044767865
## math4    0.140564617  0.764954614  0.92332515  0.951489208  0.917393931
## female   0.003009974 -0.019753464  0.01298834  0.002850811 -0.014423484
## age      0.069344630 -0.046119854 -0.09536580 -0.096909858 -0.103607528
## height  -0.049395136  0.055676186  0.04243414  0.032683477  0.040917764
##             pe3        hist4        pe4        math4       female
## treat   -0.134794987  0.21010474 -0.043975444  0.14056462  0.003009974
## read4   -0.078928737  0.79943989  0.001522679  0.76495461 -0.019753464
## read3    0.019772739  0.92717436  0.060598681  0.92332515  0.012988341
## math3    0.025802541  0.91056625  0.067988360  0.95148921  0.002850811
## hist3    0.002345894  0.96531919  0.044767865  0.91739393 -0.014423484
## pe3      1.000000000 -0.02968731  0.472814327  0.02354280  0.026070397
## hist4   -0.029687311  1.00000000  0.017628038  0.89351172 -0.017902077
## pe4      0.472814327  0.01762804  1.000000000  0.05519176 -0.013560064
## math4    0.023542799  0.89351172  0.055191759  1.00000000  0.034182940
## female   0.026070397 -0.01790208 -0.013560064  0.03418294  1.000000000
## age     -0.045765250 -0.08045850 -0.036665854 -0.11731677 -0.070468503
## height   0.046935386  0.03323997 -0.019725022  0.01223729 -0.079529570
##             age       height
## treat    0.06934463 -0.04939514
## read4   -0.04611985  0.05567619
## read3   -0.09536580  0.04243414
## math3   -0.09690986  0.03268348
## hist3   -0.10360753  0.04091776
## pe3     -0.04576525  0.04693539
```

RCT Simulation

```
namevec<-names(scores5)
#selecting variables to test
namevec<-namevec[!namevec%in%c("class","error", "treat","read4")]

#generating the syntax that will go in the lm function with a loop
allModelsList <- lapply(paste(namevec,"~treat"), as.formula)

#running all of the balance tests with a loop
allModelsResults <- lapply(allModelsList, function(x) lm(x, scores5))
```

Some of the variables you included in the balance test are problematic. Which ones?

⇒ Top Hat

RCT Simulation

`stargazer(allModelsResults[[1]],allModelsResults[[2]],allModelsResults[[3]],allModelsResults[[4]], allMode`

Table 3

	Dependent variable:				
	class_1	class_2	class_3	class_4	class_5
	(1)	(2)	(3)	(4)	(5)
treat	-0.019	0.049	-0.052	-0.052	-0.036
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
Constant	0.105^{***}	0.087^{***}	0.114^{***}	0.114^{***}	0.110^{***}
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Observations	300	300	300	300	300
R^2	0.001	0.005	0.006	0.006	0.003
Adjusted R^2	-0.003	0.002	0.003	0.003	-0.001
Residual Std. Error (df = 298)	0.301	0.300	0.300	0.300	0.301
F Statistic (df = 1; 298)	0.226	1.578	1.805	1.805	0.825

Note: *p<0.1; **p<0.05; ***p<0.01

RCT Simulation

`stargazer(allModelsResults[[6]],allModelsResults[[7]],allModelsResults[[8]],allModelsResults[[9]], allMode`

Table 4

	Dependent variable:				
	class_6	class_7	class_8	class_9	class_10
	(1)	(2)	(3)	(4)	(5)
treat	0.066*	0.015	−0.052	0.015	0.066*
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
Constant	0.082***	0.096***	0.114***	0.096***	0.082***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.020)
Observations	300	300	300	300	300
R^2	0.010	0.001	0.006	0.001	0.010
Adjusted R^2	0.006	−0.003	0.003	−0.003	0.006
Residual Std. Error (df = 298)	0.300	0.301	0.300	0.301	0.300
F Statistic (df = 1; 298)	2.866*	0.151	1.805	0.151	2.866*
Note:				*p<0.1; **p<0.05; ***p<0.01	

RCT Simulation

`stargazer(allModelsResults[[11]],allModelsResults[[12]],allModelsResults[[13]],allModelsResults[[14]], all`

Table 5

			Dependent variable:		
	read3	math3	hist3	pe3	hist4
	(1)	(2)	(3)	(4)	(5)
treat	1.329	0.934	1.038	−0.613**	4.672***
	(1.237)	(1.233)	(1.226)	(0.261)	(1.260)
Constant	78.760***	90.965***	80.911***	90.044***	79.153***
	(0.643)	(0.641)	(0.637)	(0.136)	(0.654)
Observations	300	300	300	300	300
R^2	0.004	0.002	0.002	0.018	0.044
Adjusted R^2	0.001	−0.001	−0.001	0.015	0.041
Residual Std. Error (df = 298)	9.511	9.483	9.425	2.006	9.685
F Statistic (df = 1; 298)	1.154	0.573	0.717	5.515**	13.762***

Note: *p<0.1; **p<0.05; ***p<0.01

RCT Simulation

`stargazer(allModelsResults[[16]],allModelsResults[[17]],allModelsResults[[18]],allModelsResults[[19]], al`

Table 6

	Dependent variable:				
	pe4	math4	female	age	height
	(1)	(2)	(3)	(4)	(5)
treat	−0.548	3.122**	0.003	0.043	−0.022
	(0.722)	(1.274)	(0.065)	(0.036)	(0.026)
Constant	89.953***	86.029***	0.466***	9.490***	1.295***
	(0.375)	(0.662)	(0.034)	(0.019)	(0.013)
Observations	300	300	300	300	300
R^2	0.002	0.020	0.00001	0.005	0.002
Adjusted R^2	−0.001	0.016	−0.003	0.001	−0.001
Residual Std. Error (df = 298)	5.549	9.794	0.501	0.279	0.198
F Statistic (df = 1; 298)	0.577	6.007**	0.003	1.440	0.729

Note: *p<0.1; **p<0.05; ***p<0.01

RCT Simulation

Our data seems reasonably balanced:

- a few come out as statistically significant: class_6 and class_10 at 10%,

- pe_3 at 5%.

This is the result of random chance as discussed above (we know this for certain since we modeled the data).

If you had not modeled the data, would you be concerned?

RCT Simulation

- pe_3:
    - was determined prior to treatment
    - is not generally a variable we would expect to correlated with reading scores
    - should reassure you that it is the result of random chance.
- Class_6 and Class_10 would be more concerning:
    - it might signal that some teachers were better able to get their students into the small groups
    - but the coefficients are not large, nor are they highly significant
    - should reassure you that they are the result of random chance.

If I change the seed in the simulation (try 5000), some other variables will likely be significant due to random chance.
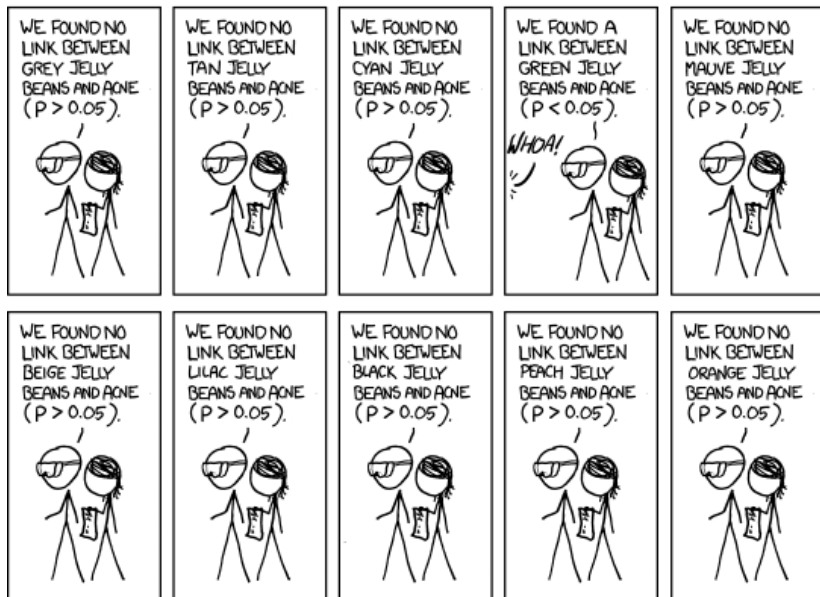
The laws of probability
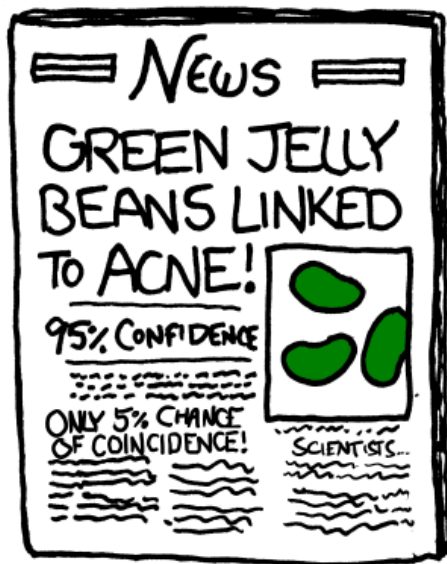
The laws of probability
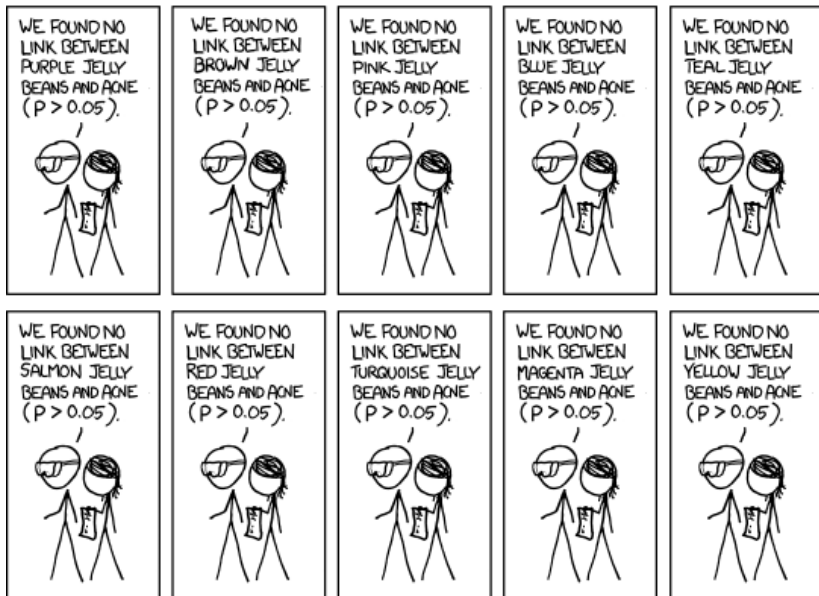
The laws of probability

The laws of probability

Controls in RCT specifications

Because the treatment was randomized, estimating

$$Y_i = \alpha + \tau D_i + \epsilon$$

▶ gives us an unbiased estimate of τ

▶ controlling for omitted variables is not necessary

That said, it is common to see specifications in RCT projects that include a vector of control variables. Why?

▶ verify that estimated coefficient does not change significantly when controls are added

▶ adding controls can make our estimated more precise and shrink the standard errors.

Controls in RCT specifications

```
rct1<-felm(read4~treat,scores5)
rct2<-felm(read4~treat+read3+female+pe3+math3+hist3,scores5)
rct3<-felm(read4~treat+read3+female+pe3+math3+hist3|class,scores5)
```

Controls in RCT specifications

```
stargazer(rct1, rct2, rct3, type="latex", header=FALSE, omit.stat = "all")
```

Table 7

	Dependent variable:		
	read4		
	(1)	(2)	(3)
treat	11.229***	10.028***	10.152***
	(1.442)	(0.824)	(0.153)
read3		0.227	0.319***
		(0.153)	(0.028)
female		−0.426	0.120
		(0.727)	(0.135)
pe3		−0.235	0.074**
		(0.181)	(0.034)
math3		0.287**	0.339***
		(0.143)	(0.026)
hist3		0.472***	0.339***
		(0.149)	(0.027)
Constant	77.150***	16.382	
	(0.750)	(16.606)	

Note: *p<0.1; **p<0.05; ***p<0.01

Controls in RCT specifications

Why does adding control variables add precision? Think about the formula for the variance/standard error of our estimator:

$$Var(\hat{\beta}_1) = \frac{\sigma^2}{SST_x(1 - R_j^2)}$$

$$se(\beta_1) = \frac{\hat{\sigma}}{\sqrt{SST_x(1 - R_j^2)}}$$

$$\hat{\sigma}^2 = \frac{1}{n - k - 1} \sum_i^n \hat{u}_i^2$$

If we include more x's in our regression,

- we can reduce $\hat{u}_1^2$, i.e. the unexplained variation in Y goes down
- $\Rightarrow se(\hat{\beta}_j)$ decreases
- $\Rightarrow \hat{\beta}$ can be estimated more precisely.

Heterogeneity

We can measure heterogeneity of the program effects for individuals with specific characteristics by interacting these characteristics with the treatment variable.

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 x_i + \beta_3 D_i \times x_i + \epsilon,$$

Example:

- x_i could be an indicator variable for being female
- β_3 gives us the differential effect of the treatment for females relative to non-females.

Simulation:

I start by searching for heterogeneity by gender using our existing simulation data.

```
rct1<-felm(read4~treat,scores5)
rcthet1<-felm(read4~treat+female+female*treat,scores5)
stargazer(rct1,rcthet1, type="latex", header=FALSE, omit.stat="ser")
```

Table 8

	Dependent variable:	
	read4	
	(1)	(2)
treat	11.229***	12.829***
	(1.442)	(1.980)
female		0.412
		(1.504)
treat:female		−3.413
		(2.893)
Constant	77.150***	76.959***
	(0.750)	(1.026)
Observations	300	300
R^2	0.169	0.173
Adjusted R^2	0.166	0.165
Note:	*p<0.1; **p<0.05; ***p<0.01	

Simulation:

I now will simulate a DGP with heterogeneous treatment effects by gender

```
#the second data generating process
nf<-20
scores5$read4het1<-(alpha+nf*scores5$treat+scores5$error+4*scores5$class_1+(-6)*scores5$class_2
                    +8*scores5$class_3+(-4)*scores5$class_4+7*scores5$class_5
                    +(-2)*scores5$class_6+5*scores5$class_7+(-10)*scores5$class_8
                    +8*scores5$class_9+4*scores5$class_10+(-20)*scores5$female*scores5$treat)

rct2<-felm(read4het1~treat,scores5)
rcthet2<-felm(read4het1~treat+female+female*treat,scores5)


#the third data generating process
nf2<-30
scores5$read4het2<-(alpha+nf2*scores5$treat+scores5$error+4*scores5$class_1+(-6)*scores5$class_2
                    +8*scores5$class_3+(-4)*scores5$class_4+7*scores5$class_5
                    +(-2)*scores5$class_6+5*scores5$class_7+(-10)*scores5$class_8
                    +8*scores5$class_9+4*scores5$class_10+(-40)*scores5$female*scores5$treat)

rct3<-felm(read4het2~treat,scores5)
rcthet3<-felm(read4het2~treat+female+female*treat,scores5)
```

Simulation:

```
stargazer(rct1,rcthet1,rct2,rcthet2,rct3,rcthet3, type="latex",header=FALSE, omit.stat="ser")
```

Table 9

	Dependent variable:					
	read4		read4het1		read4het2	
	(1)	(2)	(3)	(4)	(5)	(6)
treat	11.229***	12.829***	11.847***	22.829***	12.464***	32.829***
	(1.442)	(1.980)	(1.636)	(1.980)	(2.046)	(1.980)
female		0.412		0.412		0.412
		(1.504)		(1.504)		(1.504)
treat:female		−3.413		−23.413***		−43.413***
		(2.893)		(2.893)		(2.893)
Constant	77.150***	76.959***	77.150***	76.959***	77.150***	76.959***
	(0.750)	(1.026)	(0.850)	(1.026)	(1.063)	(1.026)
Observations	300	300	300	300	300	300
R^2	0.169	0.173	0.150	0.342	0.111	0.560
Adjusted R^2	0.166	0.165	0.147	0.336	0.108	0.556

Note: *p<0.1; **p<0.05; ***p<0.01

Simulation:

DGP 1, 2 and 3 return similar estimates of the ATE (average treatment effect).

This ATE hides important heterogenety that is quite different for DGP 1, 2 and 3:

- ▶ No heterogeneity in DGP 1

- ▶ A positive effect on non-females and no effect on females in DGP 2

- ▶ A positive effect on non-females and negative effect on females in DGP 3

Alphabet Soup!:IV's and RCT's from ITT to TOT

In RCT's, it is important to carefully define "Treatment".

Recall the blood pressure medication example: What is the "Treatment"?

- ▶ Is the treatment **taking the medication**?
- ▶ Is the treatment being **instructed to take the medication**?

With non-compliance, the estimated treatment effect will depend on how we define treatment.

If I am interested in estimating the effects of actually taking the medication:

▶ $\tau = E[Y_i(1)] - E[Y_i(0)]$ will not estimate the treatment effect.

Why?

If I am interested in estimating the effects of actually taking the medication:

- $\tau = E[Y_i(1)] - E[Y_i(0)]$ will not estimate the treatment effect.

Why?

- τ is an estimate of the **intent-to-treat (ITT)**
- it compares the outcomes of those I **intended** to give the medication to, to those I did not intend to give the medication to.
- This is the estimated effect of being **told** to take the medication.

Alphabet Soup!:IV's and RCT's from ITT to TOT

With non-complieres, the **ITT** estimate $\neq$ **TOT** estimate.

- ▶ The ITT: the effect of being in the treated group.
- ▶ The TOT: the effect of the treatment on those who actually got treated.

We can recover the TOT estimates using an IV estimation where being treated is instrumented by being in the treatment group (see the IV Module Notes).

But what do we care about? The ITT or the TOT?

But what do we care about? The ITT or the TOT?

It depends.

ITT and TOT: Contraceptive choice

▶ What is the TOT for male condoms?

▶ What is the ITT for male condoms?

▶ Which value do you care about if you are the condom manufacturer?

▶ Which value do you care about if you are a practitioner (doctor, public health advisor …)?

Table 3–2 Percentage of women experiencing an unintended pregnancy during the first year of typical use and the first year of perfect use of contraception, and the percentage continuing use at the end of the first year. United States.

Method (1)	% of Women Experiencing an Unintended Pregnancy within the First Year of Use		% of Women Continuing Use at One Year[3]
	Typical Use[1] (2)	Perfect Use[2] (3)	(4)
No method[4]	85	85	
Spermicides[5]	28	18	42
Fertility awareness-based methods	24		47
Standard Days method[6]		5	
TwoDay method[6]		4	
Ovulation method[6]		3	
Symptothermal method[6]		0.4	
Withdrawal	22	4	46
Sponge			36
Parous women	24	20	
Nulliparous women	12	9	
Condom[7]			
Female (fc)	21	5	41
Male	18	2	43
Diaphragm[8]	12	6	57
Combined pill and progestin-only pill	9	0.3	67
Evra patch	9	0.3	67
NuvaRing	9	0.3	67
Depo-Provera	6	0.2	56
Intrauterine contraceptives			
ParaGard (copper T)	0.8	0.6	78
Mirena (LNg)	0.2	0.2	80
Implanon	0.05	0.05	84
Female sterilization	0.5	0.5	100
Male sterilization	0.15	0.10	100

Emergency Contraception: Emergency contraceptive pills or insertion of a copper intra-uterine contraceptive after unprotected intercourse substantially reduces the risk of preg-nancy.[9] (See Chapter 6.).

Lactational Amenorrhea Method: LAM is a highly effective, *temporary* method of con-

- ▶ Key RCT assumption: random assignment of D
- ▶ Implicit RCT assumption: *stable unit treatment value assumption* or *SUTVA*.

SUTVA: *no spillovers.* Unit i's potential outcomes are unaffected by whether unit j is treated or untreated.

SUTVA tends to be violated whenever the treatment, D_i involves spillovers between groups, aka some type of externality.

RCT Challenges: Spillovers! SUTVA

- ▶ Treatment: receiving a measles vacination
- ▶ Outcome: getting the measles

Does SUTVA hold?

RCT Challenges: Spillovers! SUTVA

- ▶ Treatment: receiving a measles vacination
- ▶ Outcome: getting the measles

Does SUTVA hold?

- ▶ $Y_i(D_i)$ depends on all of the D's, not just D_i.
- ▶ Example: if all the units are vaccinated, except i then $Y_i(1) - Y_i(0) = 0$: no treatment effect. (Since there is no one for i to catch the measles from).

RCT Challenges: Attrition Bias

What happens to your estimates if some of the people in your experiment vanish before you can collect their endline outcome data?

- ▶ If attrition is completely random, then it is not really a problem. Your sample size will be smaller but it will not bias your results.

But what will happen if attrition is a function of treatment status?

- ▶ RCT estimates are likely biased when attrition correlates with treatment status.

Attrition Bias Simulation

To see why, lets return to our simulation of the reading group RCT.

We know that the true treatment effect is $\tau = 10$.

Suppose there are missing scores: some students did not show up on test day.

The students who do not show up are the students who know they will do poorly on the reading exam.

Attrition Bias Simulation

```
scores5$read4miss<-NA
scores5$read4miss[scores5$read4>75]<-scores5$read4[scores5$read4>75]

rctmiss<-felm(read4miss~treat,scores5)

stargazer(rctmiss, type="latex", header=FALSE, omit.stat = "all")
```

Table 10

	Dependent variable:
	read4miss
treat	6.549***
	(1.113)
Constant	84.963***
	(0.667)
Note:	*p<0.1; **p<0.05; ***p<0.01

Treatment:

▶ shifted the distribution upwards by 10 points

▶ makes the "observed" left tale longer which biases our estimate downwards

Addressing Attrition Bias

Best option:

- ▶ Minimize attrition throughout the data collection process.

If you are still faced with attrition:

- ▶ check to see if there the attrition is similar across treatment and control groups
- ▶ check if attrition correlates with observables

If you have uneven problematic attrition:

- ▶ it is sometimes possible to bound the extent of the bias by making hypothetical assumptions about who is dropping out of control and treatment.

Attrition Bias Simulation

If the principal knows that students who will score below 75 don't take the test:

▶ she can recover the true treatment effect by excluding the students who became observed because of the treatment from the estimation

Of course it is unlikely that in the real world, a researcher would know the exact model of attrition and how treatment affects attrition (ie the treatment effect in this case).

Attrition Bias Simulation

```
scores5$obsnew<-0
scores5$obsnew[scores5$treat==1 & scores5$read4>75 & scores5$read4<85]<-1

rctmiss2<-felm(read4miss~treat,scores5[scores5$obsnew==0,])


stargazer(rctmiss, rctmiss2, type="latex", header=FALSE)
```

Table 11

	Dependent variable:	
	read4miss	
	(1)	(2)
treat	6.549***	10.240***
	(1.113)	(1.102)
Constant	84.963***	84.963***
	(0.667)	(0.597)
Observations	195	177
R^2	0.152	0.330
Adjusted R^2	0.148	0.327
Residual Std. Error	7.455 (df = 193)	6.677 (df = 175)
Note:	*p<0.1; **p<0.05; ***p<0.01	

RCT advantages:

Randomized Control trials: huge advantages for causal inference.

RCTs are relatively easy to explain to policy makers, and even the general public $\Rightarrow$ important advantages when it comes to communicating research results to the wider world.

RCT have become a widely used tool in economics and the social sciences today.

The 2019 Nobel Laureat in Economics was give to Abhijit Banerjee, Esther Duflo and Michael Kremer for their role in bringing this experimental approach to economic research (particularly in development economics).

RCT limitations:

RCT's can be quite costly to conduct and the logistics of running an RCT are quite demanding.

Many very important social and economic questions where running an RCT would simply not be ethical.

- ▶ the effects of juvenile incarceration on human capital and future crime is clearly a question of first order importance but randomization would clearly be unethical.

Conceptual limitations:

- ▶ External validity
- ▶ Experimenter demand effects
- ▶ General equilibrium effects

External validity

Would you get the same results in a different context?

▶ RCT's are often conducted in a limited geographical area with a relatively small sample size.

Would you get the same results if the program were implemented on a larger scale?

▶ RCT's are often implemented with a lot more care and resources more then the large scale policy they are testing. This could change the effects.

Experimenter demand effects

Could subjects be behaving differently then they normally would because of the experimental context they are in?

- **Hawthorne effects:** the idea that individuals might modify their behavior simply because they are being observed.

- a particular concern, if they would affect the treatment and control differently.

Are subjects changeing their behavior in order to conform to what they believe the researcher expects of them?

- This is a particularly important question when experiments are incentivized and a subject could perceive that they would receive more rewards for certain types of behaviors.

Experimenter demand effects

Generally advisable to:

► Minimize the salience of the evaluation

► Ensure the experimental experience is the same for both the treated and control groups.

General equilibrium effects

Many of the policies we are interested in in economics affect variables that are not determined by a single individual's choices (eg prices).

Most RCT's are small: unlikely to affect market level variables such as prices.

If the intervention were implemented at scale, could changes in market level variables change estimated effects?